# Should the PC be Considered a Technological Revolution? Evidence from US Metropolitan Areas


Paul Beaudry
University of British Columbia and NBER

Mark Doms
Federal Reserve Bank of San Francisco

Ethan Lewis *
Dartmouth College


May 2010


*The authors would like to thank Daron Acemoglu, David Autor, David Green, Larry Katz, Thomas Lemieux, and seminar participants at UBC, LSE, UCL, UPF, the NBER Summer Institute, and the New York Federal Reserve for very helpful discussions and to Meryl Motika for research assistance.

— Abstract —

The introduction and diffusion of IT capital is widely viewed as a technological revolution. This paper uses US metropolitan area-level panel data to examine whether the links between PC adoption, educational attainment and the return to skill conform to the predictions of a model of technological revolutions. Our simple neo-classical model implies that a shift in the technological paradigm will increase the return to skill most where skill is most abundant and where the price of skill is initially low. The model also implies that the return to skill should become temporarily insensitive to increases in supply since, instead of putting downward pressure on the return, an increase in skill supply after a major technological innovation should act only to accelerate the transition to the new technology. We find that cross-metropolitan area changes between 1980 and 2000 conform closely to these and other predictions, suggesting that the era of PC diffusion may well deserve its recognition as a technological revolution.

Key Words: Biased Technological Change, Relative Wages, Education, Technology Diffusion

JEL Class.: E13, O33, J30

Paul Beaudry
Department of Economics
University of British Columbia
997-1873 East Mall
Vancouver, B.C.
Canada, V6T 1Z1
and NBER
paulbe@interchange.ubc.ca

Mark Doms
Federal Reserve
Bank of San Francisco
101 Market Street, MS 1130
San Francisco, CA 94105
mark.doms@sf.frb.org

Ethan Lewis
Economics Department
Dartmouth College
Hanover, NH 03755
ethan.g.lewis@dartmouth.edu

# 1 Introduction

The substantial rise in the return to education over the last few decades has received considerable attention from researchers and policy makers. There is now widespread consensus that changes in technology have contributed to this pattern. Despite this consensus, there remains considerable debate regarding the mechanism by which technological change has affected the return to education. The skilled-biased technological change literature includes explanations based on embodied versus disembodied technological change, and gradual versus radical change, among others. (See Saint-Paul, 2008, for an overview of the different explanations.) Many of these different explanations have been shown to be remarkably consistent with aggregate time-series observations. This state of affairs is rather unappealing given that competing explanations often differ in their policy prescriptions and their predictions for future developments.[1] Hence, it appears necessary to go beyond time series data to evaluate specific mechanisms aimed at explaining how technological change has affected the return to skill over this period.

In this paper, we ask whether the introduction of the personal computer (PC) is an example of radical change, or "technological revolution." In particular, we ask whether the recent patterns of PC adoption and the increase the return to education across US metropolitan areas conform to the predictions of a model of endogenous technology adoption describing such revolutions. The main idea in models of endogenous technology adoption (as introduced into the skill-biased technological change literature by Basu & Weil, 1998, Zeira, 1998, Caselli, 1999, and explored empirically using time series data in Beaudry & Green, 2003, 2005) is that the speed and extent of technology adoption is endogenous and reflects principles of comparative advantage.[2] According to this view, when a major new technology becomes available it is not ubiquitously or randomly adopted across space or time. Instead, it is adopted only in environments where complementary factors are cheap and abundant. Moreover, the adoption process is gradual as it is generally optimal for both the new and

[1] For example, whether raising educational attainment will reduce inequality depends on which mechanism is responsible for the increase.

[2] The model we consider can also be considered part of the "revolutions" paradigm proposed by the General Purpose Technology (GPT) literature, as it reflects the adjustment to a technological change that affects the organization of production in most sectors of the economy. (GPTs, which include computers, are essentially defined by their pervasiveness and the fact that they induce other innovations including changes to the structure of production; see Bresnahan & Trajtenberg, 1995, and Lipsey et al., 1998.) A major focus of the empirical GPT literature is the adjustment process, especially the productivity dynamics, surrounding the arrival of a new GPT. This not central in our study, but we will comment on it.

old means of production to be used simultaneously over long periods of time. Interestingly, there is a long tradition in the economic history literature that advocates the use of similar models to understand periods of major technological change.[3]

Our model generates several specific predictions which can be evaluated using metropolitan area-level data. Areas with a low relative price (or high relative supply) of skill are predicted to adopt PCs most intensively, as they have a comparative advantage in using the new skill-intensive technology. The arrival of the skill-intensive technology should cause the return to skill to increase the most in areas with initially high relative supply of skill. The model also implies that the return to skill should become temporarily insensitive to increases in supply since, instead of putting downward pressure on the return, an increase in skill supply after a major technological innovation should act only to accelerate the transition to the new technology. Finally, although the model implies that areas that use PCs most intensively experience the greatest increases in the return to skill, it is also predicts that the increase in the return to skill should not be strong enough to generate a positive association between the return and the use of PCs. As we will show, this follows directly from viewing the process of technology adoption as endogenous.

We evaluate these implications of the endogenous adoption framework using data from a sample of 217 U.S. metropolitan areas. We focus on the period 1980 to 2000, the main period of diffusion of the PC. We measure adoption of the new technology with the intensity of PC use, and measure skilled labor with college-educated labor.[4] A key challenge in bringing the model to the data is the need to credibly identify supply-driven variation in skill mix at the time PCs arrived.[5] Accordingly, we document that there are wide differences in college share across areas that originated well before PCs would have been anticipated. In particular, college share in 1940 - before even any experimental electronic computers had been developed - and high school-age enrollment rates in 1880 - before any earlier form of

[3]For instance, Goldin & Sokoloff (1984) use a similar model to explain regional outcomes across the US during the industrial revolution. They find that factor price differences in 1830 between the northern and southern U.S. states (due to crop differences) help explain the differential patterns of industrialization, and that payments to those factors rose fastest in areas where technology was adopted most aggressively. The writing of Habakkuk (1962) also reflects similar ideas. Given this previous work, it is of interest to examine whether simple neoclassical principles applied to technology adoption offer a unified way of understanding many major technological episodes, including the most recent, the IT revolution.

[4]Doms & Lewis (2006) find that the most important factor (among many examined) in understanding the large variance in PC use across cities is college share. The present paper extends Doms & Lewis (2006) by specifying and testing the mechanism that drives these differences in PC adoption.

[5]In the model, exogenous differences in skill supply are assumed to have arisen from historical forces, but we also show that the results are qualitatively the same if there are differences in amenities across areas which drive equilibrium differences in skill supply.

capital-skill complementarity took hold in the US (Goldin & Katz, 2008) - predict college share differences in 1980, on the cusp of the PC era. Using these instruments, we find that it is in areas where college-educated workers were abundant (and, we show, cheap) relative to high-school educated workers that PCs were adopted most intensely. At the same time, these areas also experienced the greatest increase in the return to education. Moreover, there was a stable downward sloping relationship between college share and the return to college in 1940 and 1980, consistent with observed skill-mix differences being supply-driven. This relationship then dissipates by 2000, as predicted by the model.[6] Despite the increase in the return, high PC adopting cities are not observed to have a higher return to education in 2000 than lesser adopters, again as the model suggests.

In our analysis, we take care to contrast the implications of our model with alternative explanations of the same facts. Although some of our observations can also be explained by alternative mechanisms, we emphasize the difficulties other explanations face in reproducing them all. For example, one potential explanation for our observation that relative wages equalized across cities between 1980 and 2000 is that trade and labor mobility across cities increased during this time period, leading to a greater increase in the return to education in cities with more educated labor. While this cannot be completely ruled out, we show that the actual movement of people and industries across cities do not provide much support for it. We also show that the most common production function specifications with PC-skill complementarity can explain some, but not all, of the cross-city patterns we document.

The remaining sections of the paper are structured as follows. In Section 2, we present a model of technology adoption and derive a set of implications regarding local level interactions between the return to education, changes in the return, and technology use. We begin by treating local supplies of skills as exogenous, then we extend the model to allow for migration across cities in response to differences in the return to skill, and thereby endogenize difference in skill mix across cities. Section 3 discusses the data and our approach to

[6] Our results may appear to conflict with more aggregate approaches, such as Katz & Murphy (1992), Krusell, Ohanian, Rios-Rull and Violante (2000) and Autor, Katz, and Kearney (2006), which suggest that there is a stable relationship between the supply of skill and its return over long periods of time, and hence the notion of technological revolution as used here may be misplaced. However, as shown in Card & Dinardo (2002) and Beaudry & Green (2005), when the aggregate US experience is split between a pre-1980 period and a post 1980 period, the relationship between the supply of skill and its return changes drastically. In particular, the relation changes in a direction consistent with the evidence presented here, as the downward effect of increased education on its return is generally insignificant in the post-1980 period. Moreover, our results are also consistent with those reported in Fortin (2006), where post-1980 state level variation in the aggregate supply of skill is not found to have any negative effects on the return to skill.

identification, and Section 4 contains our empirical results. Section 5 examines alternative explanations for our results. Section 6 offers concluding comments.

# 2 A Neoclassical Model of Technology Adoption

Consider an environment at time $t$ where firms have access to a set of technologies to produce a final good denoted by $Y_t$. The production of $Y_t$ requires inputs $Z_t$, where these inputs can be organized in different ways to produce output; each of these alternative organizations corresponds to a different technology. If we parameterize the different technologies by $\theta \in \Theta$, then the production possibilities facing a firm can be represented by:

$$F(Z_t, \theta), \qquad \theta \in \Theta_t$$

where for each $\theta \in \Theta$, the production function is assumed to satisfy constant returns to scale and concavity. In this case, a price taking firm will aim to maximize profits by solving the following problem,

$$\max_{Z_t, \theta_t} F(Z_t, \theta_t) - w_t Z_t$$

where $w_t$ is the vector of factor prices. For such an environment, standard techniques can be used to show the existence of a competitive equilibrium, where a competitive equilibrium can be defined as a set of prices, allocations and technology choices, such that, given prices, allocations and technology choices are optimal, and markets clear.

Let us now consider the situation with a set of distinct markets, indexed by $i$. Each of these markets is assumed to have access to the same set of technologies. We will begin by assuming that these markets differ in terms of the supply of at least a subset of the factors $Z$. The question we ask is how do the different markets react to a change in the set of technological choices, that is, a change in $\Theta_t$. The answer to this question depends on the nature of the change in $\Theta$. In particular, given the time period of interest, we want to examine the effects of having $\Theta$ extend to include a more skill-intensive technology relative to the pre-existing choices. To this end, we focus on the case where initially there is only one dominant technology used across all markets. This technology uses as inputs skilled labor, $S$, unskilled labor, $U$, and traditional capital, $K$. The market for skilled and unskilled labor is initially assumed to be purely local, with exogenously fixed local supplies (that may vary across markets). Later we will extend the analysis to the case where workers are mobile across markets, but where markets differ in terms of amenities and congestion. Throughout we assume that the market for $K$ is a common market, where firms from all localities can rent the capital at the rate $r^K$. Finally, for clarity of presentation and ease of comparison with the existing literature, the pre-existing technology is assumed to have the following functional form:[7]

$$F^T(K,S,U) = K^{1-\alpha}[aS^\sigma + (1-a)U^\sigma]^{\frac{\alpha}{\sigma}}, \quad 0<\alpha<1, \quad 0<a<1, \quad 0<\sigma<1$$

In this environment, the initial return to skill will differ across markets. In particular, the ratio of the market-specific skilled wage $w_i^S$ to the unskilled wage $w_i^U$ will be given by:

$$\frac{w_i^S}{w_i^U} = \frac{aS_i^{\sigma-1}}{(1-a)U_i^{\sigma-1}}$$

where $S_i$ and $U_i$ represent the quantities of skilled and unskilled labor, respectively, available in market $i$. Given the form of the production function, it is possible for the wage of skilled workers to be less than that of unskilled workers. We therefore impose Assumption 1 to insure that we focus on economies where skilled workers are paid more than unskilled.[8]

**Assumption 1:** $\frac{S_i}{U_i} < (\frac{a}{(1-a)})^{\frac{1}{1-\sigma}}$

Now suppose that at a point in time, say at $t=0$, a new technology becomes available. This technology has two characteristics that differentiate it from the traditional technology. First, it uses a different form of capital, which we denote as $PC$ capital, and $PC$ capital is assumed to be available on a common market at rental rate $r^{PC}$. Second, the new technology is assumed to be skill biased relative to the old technology in the sense that at common factor prices, the new technology uses skilled labor more intensively (i.e., has a higher ratio of $\frac{S}{U}$) than the traditional technology. These features are captured in the following functional form for the new technology:

$$F^N(PC,S,U) = PC^{1-\alpha}[bS^\sigma + (1-b)U^\sigma]^{\frac{\alpha}{\sigma}}, \qquad a<b<1$$

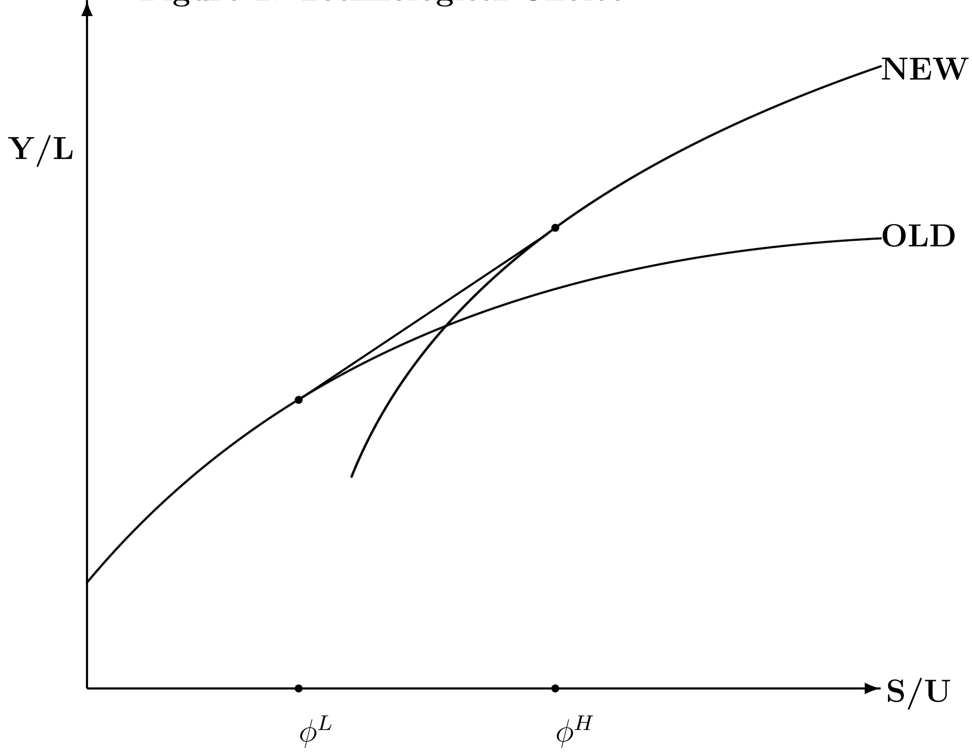
It is important to notice that the new technology does not necessarily dominate the old

[7] The analysis can be easily extended to cases where $\sigma < 0$.

[8] If we assume that skilled workers can fill jobs designated for unskilled workers, this would guarantee that unskilled workers are never paid more than skilled. For some of our results, this assumption could replace Assumption 1. However, certain results require the stronger restriction that the return to skill is positive, which is implied by Assumption 1.

technology in the sense of producing more output at any input combination. In effect, for certain rental rates of capital $r^K$ and $r^{PC}$, the old technology is more productive than the new technology when used with a small fraction of skilled workers, while the new technology is more productive (higher output per worker) when used with a high fraction of skilled workers. This property is depicted in Figure 1.

**Figure 1: Technological Choice**




It is easy to see from Figure 1 that, when faced with the choice between the new and the old technology, localities with a high ratio of skilled to unskilled workers will want to adopt the new technology, while those with a low ratio may want to maintain the old technology. We refer to the situation where there is this non-trivial choice between the two technologies as a **technological transition**. A key aspect of our analysis is to highlight the different relationship between factor prices and factor supplies during a technological transition in comparison to periods where economies do not face such choices. Note that the arrival of a new technology does not necessarily give rise to a technological transition. In our formulation, if the price of PC capital is very high, then the old technology dominates the new one regardless of the local supply of skill. As shown in Appendix A, the condition $r^{PC} > r^K \left(\frac{b^{\frac{1}{1-\sigma}} + (1-b)^{\frac{1}{1-\sigma}}}{a^{\frac{1}{1-\sigma}} + (1-a)^{\frac{1}{1-\sigma}}}\right)^{\frac{(1-\sigma)\alpha}{\sigma(1-\alpha)}}$ guarantees that no economy (which satisfies Assumption 1) will want to adopt the new technology. In contrast, if the price of PC capital is very low, then the new technology dominates the old technology. This arises when $r^{PC} < r^K(\frac{1-b}{1-a})^{(\frac{\alpha}{\sigma(1-\alpha)})}$, in which case the new technology would be adopted by all profit-maximizing firms regardless of local market conditions. Hence, a technological transition corresponds to a temporary period that arises after the initial development of the technology and lasts until other developments allow the new technology to become dominant. Since we want to focus on the implications of a technological transition, we pursue our analysis under Assumption 2.

**Assumption 2:** The rental price of PC satisfies $r^K(\frac{1-b}{1-a})^{(\frac{\alpha}{\sigma(1-\alpha)})} < r^{PC} < r^K(\frac{b^{\frac{1}{1-\sigma}}+(1-b)^{\frac{1}{1-\sigma}}}{a^{\frac{1}{1-\sigma}}+(1-a)^{\frac{1}{1-\sigma}}})^{\frac{(1-\sigma)\alpha}{\sigma(1-\alpha)}}$

Under Assumption 2, some localities may find it optimal to adopt the new technology while others will not. However, it is not the case that localities will either fully adopt the old or new technology. Instead, the adoption decision is characterized by three regions delimited by skilled to unskilled labor ratios, with the middle region involving localities where the new and old technology co-exist. As shown in Appendix A, there exist critical values of skill ratios $\phi^L$ and $\phi^H$ ($0 < \phi^L < \phi^H$) such that if a locality is characterized by $\frac{S_i}{U_i} < \phi^L$, then it maintains the old technology. If $\frac{S_i}{U_i} > \phi^H > \phi^L$, then the locality switches completely to the new technology. Finally if $\phi^L < \frac{S_i}{U_i} < \phi^H$, then both technologies co-exist in a competitive equilibrium, and the fraction produced with the new technology will be an increasing function of $\frac{S_i}{U_i}$. We refer to a locality where the technologies co-exit as a locality that is **experiencing the technological transition**.

Proposition 1 addresses the link between $PC$ use and local skill supply. Since $PC$ capital is used intensively in the new technology, it follows that the quantity of PCs per worker used in a locality is a monotonically increasing function of the ratio of skilled to unskilled workers.[9] This forms the basis of Proposition 1.

[9] We have presented what we believe to be the simplest model that delivers the propositions which we investigate empirically. However, on some dimensions, it is certainly too simplistic. For example, the particular formulation implies that adding more unskilled workers to a labor market while keeping the number of skilled workers constant would lead to a decrease in the number of PCs. This implication is not robust as it results from the extreme assumption that there is no possibility of using PCs in the traditional technology. A slightly generalized formulation, which reverses this implication, is one where the traditional technology uses both structures and equipment. Assume that at time $t = 0$, the PC becomes available. The PC can then be used either as a substitute for equipment in the traditional technology, or it can be used in a new form of organization which is both skill-biased and uses PC more intensely than the traditional technology (in the sense that, as given factor prices, it uses a greater number of PCs per worker). This latter form of work organization is what we envision as the new technology. It can be easily verified that this alternative formulation is consistent with all of the propositions presented in the paper, but it does not imply that the number of PCs used would decrease in response to an increase in unskilled labor.

**Proposition 1:** After the arrival of a PC-based, skill-biased technology, the ratio of PCs per worker will be an increasing function of a locality's ratio of skilled to unskilled workers.

**Proof:** The proofs of all propositions and corollaries are given in Appendix A.

Proposition 1 indicates that skill-biased technologies are adopted most aggressively by localities in which skill is relatively abundant, and therefore the observable aspects of the technology – such as $PC$ capital — are most prevalent in localities with more skill. This implication is the focus of Doms & Lewis (2006). Here, we want to go further and derive a set of additional implications in order to more closely examine the relevance of a biased technology adoption model for understanding differences in outcomes across localities. To this end, we extend Proposition 1 and derive a corollary that captures the incentive mechanism that leads to the different adoption decisions. Note that from an individual firm's perspective, the differential adoption decisions across localities must reflect different incentives induced by factor *prices*. In localities with initially high ratios of skilled to unskilled labor – which, in the model, is synonymous with a high relative supply of skilled labor – the relative price of skilled labor is initially low (prior to the availability of the new technology), favoring the adoption of a technology which uses skill intensively. This is Corollary 1.

**Corollary 1:** The ratio of PCs per worker is a decreasing function of a locality's initial ratio of skilled to unskilled wages.

Proposition 1 and Corollary 1 focus on the effects of local market conditions on adoption decisions. We now want to examine the more intriguing aspect of the technological transition: how the arrival of the new technology affects the relationship between factor prices and supply. In particular, we first want to emphasize how changes in the return to skill, as expressed by the change in $\log(\frac{w_i^S}{w_i^U})$, varies across localities faced with the same technology options. This is captured in Proposition 2 and Corollary 2.

**Proposition 2:** The arrival of the skill-biased technology causes the return to skill to increase most in localities where skill is abundant.

The content of Proposition 2 can be obtained by deriving the relationship between the return to skill and the supply of skill before and after the arrival of the new technology, and taking the difference between the two. This relationship is expressed analytically below (and graphically in Figure 2, below). For localities with very low initial supply of skilled workers, relative wages do not change since the new technology is not adopted. Localities where $\phi^H < \frac{S_i}{U_i}$ experience the largest increase in the return to skill since they switch entirely to the new technology which causes an increase in the demand for skill. Finally, for localities in the partial adoption region $\phi^L < \frac{S_i}{U_i} < \phi^H$, the increase in the return to skill is strictly increasing in the relative supply of skill since the endogenously induced demand for skill is increasing with skill.

$$\begin{array}{llll}
\Delta \ln \frac{w_i^s}{w_i^U} = & 0 & if & \frac{S_i}{U_i} \leq \phi^L \\
\Delta \ln \frac{w_i^s}{w_i^U} = & (1-\sigma)[\log \frac{S_i}{U_i} - \log \phi^L] & if & \phi^L < \frac{S_i}{U_i} \leq \phi^H \\
\Delta \ln \frac{w_i^s}{w_i^U} = & (1-\sigma)[\log \phi^H - \log \phi^L] & if & \phi^H < \frac{S_i}{U_i}
\end{array}$$

Proposition 2 expresses how the arrival of the new technology induces a positive association between the supply of skill and changes in the return to skills. However, the proposition bypasses the channel through which this arises. Corollary 2 addresses this issue by combining Propositions 1 and 2 to highlight how it is the adoption of the PC-intensive technology that leads to increases in the returns to skill.

**Corollary 2:** The return to skill increases the most in localities which choose to adopt PCs most intensively.[10]

At first pass, Proposition 2 may appear counterintuitive since it predicts an increase in return to skill where supply is most abundant. However, this does not imply that the arrival of the new technology can cause the level of the return to skill to be positively related to supply. In fact, as stated in Proposition 3, even after the introduction of the skill-biased technology, the return to skill must remain a weakly decreasing function of the supply of skill. Note that it is possible for the arrival of the skill-biased technology to cause the disappearance of a negative relationship between return and supply if localities are concentrated in the technology-mixing zone ($\phi^L < \frac{S_i}{U_i} < \phi^H$), but the relationship between the return to and the supply of skill cannot be positive even after the introduction of the skill-biased technology.

[10] Any approach taken to evaluate Corollary 2 must acknowledge the endogeneity between PCs and returns to skill. Corollary 2 implies that it is the adoption of PCs, induced by differences in initial supply of skill (or initial returns to skill), that causes the return to skill to increase.

**Proposition 3:** The arrival of the skill-biased technology cannot induce a positive association between the return to skill and the supply of skill.

The content of Propositions 2 and 3 can be easily inferred from Figure 1. Because the return to skill is captured by the slope of the production function, one can notice that the slope of the outer envelop is weakly decreasing in the fraction of skilled workers. This is the content of Proposition 3. In contrast, if we consider the change in the return to skill induced by the new technology for an initial supply in the region ($\phi^L < \frac{S_i}{U_i} < \phi^H$), we see that the increase in the slope is larger for initially higher levels of supply. The reason is that the return to skill was initially more depressed in the higher supply localities and therefore the new technology allows for greater induced demand for skill in such areas. The content of Proposition 3 is depicted in Figure 2. In this figure, we see that the availability of the new technology alters the relationship between returns to skill and supply. However, the slope of the new relationship is nowhere positive. Note that in the region $\phi^L < \frac{S_i}{U_i} < \phi^H$, the slope of the relationship is zero since the technological choice allows the reallocation of additional skill between the two technologies without affecting the return to skill.[11]

Now that we have examined the effects of skill supply on both technology adoption and the return to skill, we can combine the two to obtain Corollary 3.

**Corollary 3:** The return to skill will not be larger in localities with more intensive PC use.

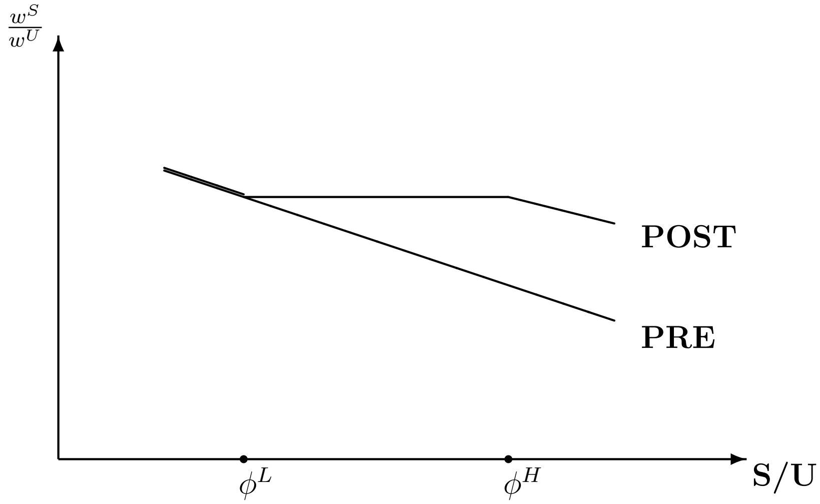
Corollary 3 indicates that although PC adoption and increases in the returns to skill should go hand-in-hand (as stated in Corollary 2), such positive co-movement cannot induce an outcome where the return to skill is higher in localities with high PC-intensity than localities with low PC-intensity. This follows directly from the endogenous PC adoption framework. To be more precise, PCs are adopted more aggressively in one locality versus another only because the cost of skill is lower. Therefore, PC capital cannot be more intensely used in a locality with a higher cost of skill. By contrast, if the adoption and subsequent use of PCs were viewed as an exogenous phenomena (as is the case in many papers), then it would be natural to expect to find a positive association with PC use and returns to skill.[12] Hence, this prediction nicely illustrates how a model of endogenous technology adoption differs from more conventional models with exogenous technological change.[13]

[11] This mechanism is identical to the factor price equalization zones in international trade theory.

[12] Assuming PCs are skill-intensive. For example, the fact that the return to education is larger for computer users (Krueger, 1993) was once considered crucial evidence that computers and skill were complements.

[13] The implication of endogenous adoption stated in the previous propositions and corollaries would be

**Figure 2: Effect of Relative Supply on the Return to Skill**




From Figure 2, one can also observe that an increase in skill has a different effect on its return prior to the arrival of the new technology than during the technological transition. As noted in Proposition 4, an increase in supply has less of an effect on the return to skill during the transition, since for localities experiencing the transition, an increase in supply causes them to allocate a greater fraction of the workforce to the new technology, thereby allowing the return to skill to stay constant as supply increases.

**Proposition 4:** A local increase in the supply of skill has a less negative effect on the return to skill during a technological transition. In particular, for localities experiencing the transition, an increase in skill only accentuates the adoption of the new technology and does not decrease the return to skill.

## 2.1 Allowing for Labor Mobility

We have so far emphasized the role of differences in skill mix at the city level on technology adoption and on the subsequent induced change in wages. As we start to think about how to evaluate this model empirically (Section 3) it bears emphasizing that empirically relevant skill-mix differences are *supply*-driven.[14] In our model so far, labor supply is assumed

modified if the adoption process involved externalities. For example, suppose there existed a network type externality associated with the adoption of the new technology, so that as more local production was done with the new technology, the productive performance of the new technology increased. In such a case, it is possible that the return to skill would be positively related to PC-intensity. Also, the return to skill could be, at least over a range, an increasing function of the supply of skill.

[14]Whether such skill supply differences can actually be identified is another question, and one which we address in Section 3 where we talk about our empirical strategy.

exogenous and fixed, but in the real world labor supply is mobile across markets. So it is also important to examine whether the above results still hold if we relax this assumption and allow skill mix to endogenously react to changes in relative wages.

To this end, we extend our model to allow workers to move between localities in response to differences in expected utility. However, even if workers are perfectly mobile across localities this will not necessarily lead to equal wages since localities may have other attributes that affect their attractiveness.[15] Moreover, these local characteristics will generally be valued by skilled and unskilled workers differently, which will cause skill mix to vary geographically.

To capture the notion that some localities have a comparative advantage in attracting a more skilled labor force, let us consider the situation where there are two types of housing in each city, one that is attractive to skilled workers and one that is attractive to unskilled workers; and let us assume that localities differ in their capacity to supply one type of housing versus the other.[16] To be more precise, let preferences of workers depend on their consumption of a traded good, c, and their consumption of housing services $h^s$ or $h^u$. For a skilled worker, preferences are captured by the utility function $\log(c) + \upsilon \log(h^s)$, and for an unskilled worker preferences are given by $\log(c) + \upsilon \log(h^u)$. The consumption good is assumed to be tradeable across cities and to be produced using the technologies discussed in the previous section; that is, technological change will take the form of a change in the technology set available to produce $c$. The two types of housing are assumed to be locally produced with the production possibility set of a locality given by:

$$[d_i(H_i^s)^\psi + (H_i^u)^\psi]^{\frac{1}{\psi}} = E_i \qquad \psi > 1$$

where $H_i^s$ is the total supply of housing directed to skilled workers in market $i$, $H_i^u$ is the total supply of housing directed to unskilled workers in market $i$, and $E_i$ is the land endowment of market $i$. For simplicity, we assume that land is the only input used to produce the non-tradeable good called housing. Note that the parameter $d_i$ in the production possibility set governs a city's comparative advantage for providing housing that is attractive to skilled versus unskilled workers. A city with a low value of $d$ has a comparative advantage

[15] As discussed in Black, Kolesnikova and Taylor (2004), amenities can help explain differences in skill mix, differences in the return to skill, and differences in housing prices across cities.

[16] The assumption that housing preferences are directly associated to skill is a shorthand way of of capturing the differential demands of the two groups for local amenities. An alternative, but more complicated, approach is to endogenously derive the different demands through income effects.

in supplying housing that is desirable to skilled workers. Given a competitive market for housing, the production possibility set implies that the relative prices for housing services will satisfy:

$$\frac{p_i^s}{p_i^u} = d_i(\frac{H_i^s}{H_i^u})^{\psi-1} = d_i(\frac{S_i h_i^s}{U_i h_i^u})^{\psi-1}$$

where $S_i h_i^s$ represents the total demand for housing by skilled workers in market $i$, and $U_i h_i^u$ represents the total demand for housing by the unskilled.

Workers take wages and house prices for each locality as given when deciding where to locate and how much housing to consume. Utility maximization on the part of workers implies that the indirect utility in locality $i$ is $(1+\upsilon)\log(w_i^s) - \upsilon \log p_i^s$ for a skilled worker and is $(1+\upsilon)\log(w_i^u) - \upsilon \log p_i^u$ for an unskilled worker, where $p_i^s$ and $p_i^u$ are the prices of housing in locality $i$. The associated quantity of housing demanded in locality $i$ is $h_i^s = \frac{\upsilon w_i^s}{(1+\upsilon)p_i^s}$ for a skilled individual and $h_i^u = \frac{\upsilon w_i^u}{(1+\upsilon)p_i^u}$ for an unskilled individual. Since workers will migrate to ensure equal utility across localities, relative skill premiums will be proportional to the difference in relative housing prices between any two localities $i$ and $j$, as expressed below:

$$\log(\frac{w_i^s}{w_i^u}) - \log(\frac{w_j^s}{w_j^u}) = \upsilon[\log(\frac{p_i^s}{p_i^u}) - \log(\frac{p_j^s}{p_j^u})]$$

The above relationships can be combined to provide a simple expression for the relative supply of skill as a function of the relative wages between two localities $i$ and $j$. This is given by the equation below where we can see that a locality will attract more skilled workers if it either has a comparative advantage in supplying desirable housing (low $d_i$) or if the skill premium is high.

$$\log(\frac{S_i}{U_i}) - \log(\frac{S_j}{U_j}) = \frac{\psi}{\psi-1}[\log(\frac{d_j}{d_i}) + (\frac{1+\upsilon}{\upsilon})(\log(\frac{w_i^s}{w_i^u}) - \log(\frac{w_j^s}{w_j^u}))]$$

If we now consider the situation before the technological transition, we know that demand for skills by firms producing the tradeable consumption good in locality $i$ is given by $\frac{aS_i^{\sigma-1}}{(1-a)U_i^{\sigma-1}}$. Combining the relative demand relation with the supply relationship allows us to express relative wages and relative supplies as a function of the comparative advantage parameters $d_i$. It is easy to verify that is the case. The mobility of workers will lead a locality with a low $d_i$ to have both a greater relative supply of skilled workers, and a lower return to skill. This represents the situation prior to the arrival of a new technology. The question we now want to address is how allowing for the mobility of workers across cities affects our results regarding technological transition. This is stated in Proposition 5.

**Proposition 5:** When skill supply across localities adjusts to equate utility, a technological transition will still lead to (1) $PC$s per worker being higher where relative skill supply is initially highest, (2) the return to skill to increase most where skill supply is initially highest. However a technological transition will not lead to a situation where the return to skill is positively associated with either higher levels of $PC$ per capita or the supply of skill.

Proposition 5 indicates that the results from the previous section, where workers were assumed immobile, can be extended to allow for worker mobility.[17] The intuition for why these results extend easily to a case with mobility can be inferred by adding to Figure 2 a positively sloped supply curve for skill. In this modified figure, localities with low $d$'s would have relative supply curves for skill that are translated to the right relative to those for localities with a high value of $d$. The technological transition causes the relative demand for skill to pivot upward, moving along the supply for skill schedules where, among any two localities experiencing the transition, there would a greater increase in return to skill where initial supply is highest. This is not to say that allowing worker mobility has no effect on the equilibrium outcome. Clearly, it would have quantitative effects. However, Proposition 5 indicates that it would not change the qualitative implications of Propositions 1 through 3. Interestingly, when worker mobility is modeled along the lines presented here, a technological transition would be a period where relative skill supplies diverge across localities, with high skill cities becoming even more highly skilled. This arises because the arrival of the new technology counters the decreasing returns to skill and thereby favors in-migration of skilled workers.[18] Such an outcome could at first pass be mistaken for the emergence of some form of positive externality associated with skill. However, the current model offers an

[17] The only Proposition which can not be directly extended to the case of worker mobility is Proposition 4 since it is no longer meaningful to talk about an exogenous increase in the supply of skill.

[18] The evidence on whether or not there has been a divergence across cities in the relative supply of skill over 1980-2000 is somewhat mixed. If one follows Berry & Glaeser (2005) and defines skilled workers as workers with at least a 4 year college degree, then there appears to have been divergence. In contrast, if one follows Katz & Murphy (1992) and also include individuals with a non-BA post-secondary degree in the skilled group, then the data does not reveal any divergence. Further, these results are also sensitive to whether relative skill supplies are measured in logs or in levels. Since the main focus of this paper is not whether a technological transition causes divergence in the geographical distribution of skill, we do not pursue this point in the empirical section.

explanation for such a divergence by the arrival of a new technology choice which does not involve any externalities.

## 2.2 Relationship to Other Models

There are many antecedents to this model which feature endogenous technical change, including Caselli (1999), Zeira (1998) and Basu & Weil (1998). One that comes very close to the present paper is Goldin & Sokoloff (1984). They modeled the adoption of large scale manufacturing in the nineteenth century, which they argued was unskilled-intensive and so was adopted in the location where unskilled labor was available (the North, where unskilled labor was not tied up agriculture). Rather than modeling the arrival of a new sector, however, we are modeling the arrival of a new technology within a sector, or really, a new technology which affects many sectors in roughly the same way. The model here is also similar to Beaudry & Green (2003, 2005). The original BG papers rely on aggregate time-series variation, while the present paper's evaluation relies only on variation within the US, allowing for potentially more convincing identification (Section 3.1) and a better chance to distinguish it from other models with overlapping aggregate implications (Saint Paul, 2008). It is also fruitful to evaluate the model in a context – local labor markets – outside the one for which it was originally conceived, with appropriate modifications.[19]

Our model also has much overlap with papers which model computerization as an example of the arrival of a new "General Purpose Technology" (GPT), e.g., Bresnahan & Trajtenberg (1995). GPTs are generally characterized as having widespread effects, gradual adoption (the reasons for which include the fact that the technology is improving, and new uses for it are being developed), and leading to the reorganization of production, which is very similar to how we describe PCs.[20] The studies in this literature mainly focus on the adjustment to a new GPT, especially the productivity dynamics. A key feature is a period in which productivity growth slows as resources are diverted to researching how to benefit from the GPT (e.g., Helpman & Trajtenberg, 1998), a view for which there is empirical support (e.g.,

[19]In particular, capital supply is elastic in this local labor market version, and we have allowed for labor mobility (section 2.1).

[20]A nice historical example, described in Lipsey et al. (1998), is how electricity allowed reorganization of the manufacturing shop, allowing subsequent efficiency gains. Essentially, as a result of electricity machines could be moved away from a central power source, and instead moved into order of processing. Our model abstracts from such organizational change details, other than to say effective use of the new technology involves greater skill intensity.

Basu & Fernald, 2008). These sorts of dynamics are not the focus of our study.[21] However, will comment on the average wage dynamics that we find in our data in section 4.3.

GPT models also sometimes attempt to account for changes in the skill premium, working again through the temporary diversion of skilled labor to increased research activity (e.g., Helpman & Trajtenberg, 1998). A related idea, going back to Nelson & Phelps (1966), is that skilled workers have a comparative advantage in implementing or adapting to new technology, and so the arrival of new technology temporarily raises demand for skilled labor (Galor & Moav, 2000; Greenwood & Yorukoglu, 1997; Bartel & Lichtenberg, 1987). This implies any increase in the skill premium is temporary – it goes away after the implementation phase is over. In section 4.2 we look for evidence of recent declines in the skill premium.

The model in the present paper also encompasses the idea that PCs and skills are complements, and so is related to studies of "skill-biased technical change" (many papers, but, for example Katz & Murphy, 1992; Krusell et al., 2000). In the SBTC view, the fall in the price of PCs drives up relative demand for skills.[22] However, at least standard "CES" ways of modeling PC-skill complementarity appear unable, in general, to capture all of our model's predictions – in particular the change in the relationship between the supply of and the return to skill – something we discuss in section 4.3.

# 3 Data and Estimation Strategy

Section 2 highlighted several implications of viewing technological adoption as driven by principles of comparative advantage. Our goal now is to examine whether metropolitan area-level outcomes – hereafter, "city-level" for short – observed mainly over the 1980-2000 period exhibit the patterns implied by such a model. We choose to focus on this period for several reasons. First, this is a period often considered one of technological revolution due to astounding technical progress and diffusion of information technology. Hence, it is a perfect candidate period to see whether a neo-classical model of technological adoption is relevant. Second, it is a period in which the return to education increased substantially, and skill-biased technological change is often considered to be one of the reasons behind

[21] Instead, our model might be best be described as focusing on two points in time: before the GPT has had much impact, and in the middle of the adjustment process.

[22] Goldin & Katz (2008) model computers as just another form of capital. Their view is that a steady fall in the price of capital that long predated the arrival of computers has steadily driven up skill demand; what changed recently is that relative skill supply has grown more slowly.

this increase. Therefore, it is particularly relevant to examine whether this period is best characterized as reflecting the effects of exogenous technological change (in line with much of the literature) or whether instead it reflects an endogenous choice of technique.[23]

The city-level data we use can be roughly divided into two categories: technology and demographic. The technology data is derived from establishment-level surveys on technology use carried out by a private marketing company (Harte-Hanks) and is described in more detail in Doms & Lewis (2006). About 160,000 establishment-level observations are used to compute the adjusted average PCs per employee in each city in our sample (217 cities).[24] In computing adjusted PC intensity, we control for the three-digit SIC industry interacted with 8 establishment size classes, for a total of over 1,800 interactions.[25]

We focus on PCs instead of other IT technologies for several reasons. First, businesses spent about 90 percent more money on PCs during the 1990s than on other types of computers. Also, spending on PCs is likely correlated with other information technology spending, such as spending on software, computer networking equipment, printers, et cetera. Finally, we were able to obtain consistent measures of PCs over this period.[26]

To get a sense of the wide variation in adjusted PC use across metropolitan areas, Figure 3 shows a scatter plot of our city-level PC measure for 2000 against a similarly constructed measure for 1990.[27] The figure shows quite a bit of variation across areas in PC use (standard deviation is 0.07 in 2000) which appears to be relatively stable, or even slightly diverging. To pick two cities near the extremes, in 1990, the mean establishment in San Francisco had .12 more PCs per employee than the mean establishment in Scranton, PA after controlling

[23] Note that it could be possible that the extent of bias of technical change is endogenous at the national level, but not at the city level since markets across cities are well integrated. In such a case, our approach of focusing on city-level outcomes would not identify elements of endogenous technological change. In other words, our empirical work evaluates the joint hypothesis that technological adoption is a phenomena that reacts to market conditions and that the labor markets in cities across the US are not perfectly integrated.

[24] To increase the precision of our city-level measures, the PC intensity measure we describe as "2000" combines 2000 and 2002 data. Doms & Lewis (2006) define "city" primarily as consolidated metropolitan statistical areas (CMSAs). The logic was to derive city definitions that corresponded to the idea of a local labor markets. Our results throughout the paper are insensitive to how we define labor markets in especially large, contiguous areas, such as in and around New York City.

[25] The SIC-size interaction allows for the possibility that, for instance, large banks perform different operations than small banks. As described in Doms & Lewis (2006), our city-level measures of PC intensity are strongly correlated with other measures that control for 4-digit SIC and for measures that also control for the firm to which an establishment belongs.

[26] Other measures of information technology were examined in Doms & Lewis (2006), including more refined measures of PC intensity. The results in Doms & Lewis (2006) were robust to the choice of technology measure.

[27] The "1990" adjusted means are constructed by stacking 1990 and 1992 Harte-Hanks surveys.

for industry and establishment size differences across the two cities. In 2000, the difference in PC intensity between the Bay Area and Scranton increased to .16. More broadly, the linear slope coefficient on 1990 versus 2000 is slightly over one, with an R-squared of 0.5. This figure is also a reminder that most of the adoption of personal computers took place in the 1990s: in 1990, the median area had one PC for every *six* workers; by 2000, this had risen to one PC for every *two* workers. The adoption of PCs has likely continued to increase in more recent years (likely at a somewhat slower pace), though unfortunately we do not have the data to demonstrate this directly.[28] As a result, we will only briefly examine more recent outcomes, and only for wage outcomes.

Data on wages and skill mix comes from the decennial US censuses, specifically the public-use micro-data files for 1880, 1940, 1950, 1980, and 2000.[29] To get a more recent year we use the stacked 2006-8 American Community Surveys (Ruggles et al., 2009) which are surveys similar to the 2000 Census (but with fewer observations, which is why we stack multiple years). Using these data, we define skill supply using "college equivalents," defined as workers who have a least a four year college degree plus one-half of those with at least some college education. We empirically implement relative skill supply, $(\frac{S}{U})$, as the the log ratio of college equivalents to non-college equivalents. Measures similar to this one have often been used in research examining the effects of skill-biased technological change, such as Katz & Murphy (1992), Autor, Levy and Murnane (2003), and Card & DiNardo (2002).[30]

Figure 4 shows a scatter plot of our skill measure in 1980 and 2000. As the figure shows, the measure of skill varies greatly across cities (standard deviation about 0.3 in both years), but as with the PC intensity data, there is great persistence in the skill mix, something we take advantage of in our IV strategies described in the next section. The skill mix in 1980 explains over 80 percent of the variance in 2000.

In most cases, our measure of relative wages is computed using the average log hourly earnings of people who report completing exactly 12 years of education and people who report completing exactly four years of post high school education.[31] We do not use the raw (log)

[28]The share of workers using a computer at work rose roughly linearly 1993 through 2003, the last year that the Current Population Survey tracked computer use at work, according to Valletta (2006).

[29]We also make occasional reference to calculations using the 1970 and 1990 censuses.

[30]Our sample of workers includes both men and women aged 16-65 with at least one year of potential work experience, employed, not living in group quarters. The wage sample is further restricted to those with wages between \$2 and \$200 in 1999 dollars. See Appendix B.

[31]As a robustness check, we compute the returns to college using the sample of all workers with at least 11 years of education in Table 3b, and estimate returns to college equivalent status.

wage ratio (though our results are robust to doing so) but rather the wage ratio adjusted for a fourth-degree polynomial in potential work experience, and dummies for female, immigrant, age, foreign-born, and married. Appendix B has more details on the wage construction. In some cases we have also attempted to correct returns for self-selection into particular labor markets using Dahl's (2002) selection correction methodology. The details of our implementation of Dahl are similar to Beaudry, Green, and Sand (2007) and are described in Appendix B.

A key outcome is the rise in this wage gap between 1980 and 2000, and we weight regressions to be efficiently estimated for this outcome (though this turns out to make little difference in the estimates or the standard errors).[32] To be consistent, we use these weights throughout.

To reduce omitted variables bias (discussed further in the next section) we control for several city-level demographic measures that we label as "city controls": the log of the size of the area's labor force and the percent of the area's workforce which is African American, female, Mexican-born, and U.S. citizens in 1980. Additionally, we sometimes include "industry controls" which reflect a city's employment mix across 12 industry groups in 1980.[33]

## 3.1 Estimation Strategy

In the next section we will examine whether the patterns of PC adoption, return to skill, and education levels in the city-level data described in the previous section conform to the predictions of the model outlined in Section 2. The model highlights how forces of comparative advantage would cause cities with initial differences in relative skill supply to react differently to a change in the technological paradigm. In the model, initial differences in skill mix were assumed to be exogenous (supply-driven), coming from factors such as historical events unrelated to the new technological opportunities. This is our maintained assumption for the empirical analysis. Since this assumption may not hold, we subject it to

[32]In particular, regressions are weighted by $(\frac{1}{N_{12,1980}} + \frac{1}{N_{16,1980}} + \frac{1}{N_{12,2000}} + \frac{1}{N_{16,2000}})^{-0.5}$, where $N_{16}$ and $N_{12}$ represent the raw number of wage-reporting census respondents in each area that have exactly 16 and 12 years, respectively, of education in the corresponding census. This weight is highly related to city size: it has a correlation of 0.9 with 2000 employment and labor force.

[33]These give the share of employment in industry categories which correspond roughly to one-digit SIC: agriculture & mining, construction, non-durable manufacturing, durable manufacturing, transportation & utilities, wholesale, retail, "FIRE," business & repair services, other low-skill services, entertainment, and professional services (public sector excluded category). Note that in the PC regressions, these industry mix controls are *on top of* the detailed industry adjustment already performed on the dependent variable (three-digit SIC x establishment size). The industry mix controls therefore capture any additional indirect or "spillover" effects of industry mix in PC regressions.

several falsification exercises using other sources of historical variation in skill mix.

A general estimation equation helps highlight the potential endogeneity issues:

$$\Delta Y_{i,2000-1980} = \gamma_0 + \gamma_1 \ln(\frac{S}{U})_{i,1980} + \epsilon_{i,2000-1980} \qquad (1)$$

where $\Delta Y_{i,2000-1980}$ represents the change in outcome "$Y$" (PC adoption or change in the college-high school ln wage gap) in city $i$ between 1980 and 2000 and $\ln(\frac{S}{U})_{i,1980}$ represents the initial educational mix of the area.[34] Estimates of (1) will be used to evaluate the model's prediction about the impact of initial skill mix on $\Delta Y$. However, other unobserved factors, $\epsilon_{i,2000-1980}$, may also affect $\Delta Y$ and may also be correlated with $\ln(\frac{S}{U})$, biasing OLS estimates of (1). We can imagine at least two broad sources of bias. First, it may the case that workers can partially predict $\epsilon_{i,2000-1980}$ and move in anticipation. We will refer to this source of bias loosely as "reverse causality." In particular, $\epsilon_{i,2000-1980}$ may include city-specific long-run trends in the relative demand for skilled labor which tend to both raise wages and attract skilled labor.[35] Some such forces could simultaneously raise demand for PCs in the period of our study.[36] Note that this type of bias suggests an empirical test to potentially falsify the assumption that 1980 differences in skill mix are supply driven: 1980 skill mix should be uncorrelated with prior changes in similar outcomes. In particular, we will test whether 1980 skill mix is related to trends in the return to skill *prior* to 1980; if it is, it will falsify the assumption that 1980 skill mix can be treated as exogenous.

Second, there may be a stock of some other factor input, $z$, that is both correlated with the stock of skilled workers – it may be complementary to skill – and increases the return to skill and PC adoption between 1980 and 2000. Modifying (1), we have:

$$\Delta Y_{i,2000-1980} = \gamma_0 + \gamma_1 \ln(\frac{S}{U})_{i,1980} + \gamma_2 z_{i,1980} + \mu_{i,2000-1980} \qquad (1')$$

The error term from (1) has been expanded as $\epsilon_{i,2000-1980} = \gamma_2 z_{i,1980} + \mu_{i,2000-1980}$ (where $\mu_{i,2000-1980}$ is defined to be orthogonal to $\ln(\frac{S}{U})_{i,1980}$).

[34] (1) suppresses the fact that the coefficients, $\gamma$, and the error are outcome specific. Specifying this, the right-hand side would be $\gamma_0^y + \gamma_1^y \ln(\frac{S}{U})_{i,1980} + \epsilon^y_{i,2000-1980}$ for each outcome $y$.

[35] Feeding this, there is evidence that college graduates are more responsive than non-college workers to labor market conditions in external markets (Wozniak, 2006).

[36] Note that they would not *necessarily* do so, since they would also tend to make skilled labor more expensive.

This source of bias is more challenging to address because $z$ could be many things: it could include stocks of pre-PC equipment, unobserved worker quality or skills, or the local business climate (taxes, regulations, etc.), for example. For concreteness, we will refer to cities with a high level of this unobserved input as "innovative" cities. We include controls (described in the previous section) to try to partially alleviate such concerns. However, the "innovative city" factor may be difficult to measure directly.

To address this, we start by using a long lag of $\ln(\frac{S}{U})$ as an instrument for 1980. In particular, we use 1940, the earliest year in which educational attainment is measured in nationally representative data.[37] This is a valid instrumental variable strategy if $\ln(\frac{S}{U})_{i,1940}$ is uncorrelated with $z_{i,1980}$. While it is plausible that education differences across cities in 1940 did not anticipate the computer revolution 40 years later – in which case this instrument helps address the "reverse causality" sort of bias – it might still be the case that there is some unobserved factor $z$ which was also related to skill mix in 1940 and is also highly autocorrelated. As a result, $\ln(\frac{S}{U})_{i,1940}$ may pick up the effects of $z_{i,1940}$.

However, it turns out there is a way to test for this sort of bias if (we think plausibly) over this 40 year period *either* the autocorrelation of $z$ and $\ln(\frac{S}{U})$ differ, or the relationship between them changes (or both). Assuming for the moment this is so (reconsidered below) we can evaluate this source of bias by breaking up the variation in $\ln(\frac{S}{U})$. In particular, if we decompose 1980 education mix as the 1940 skill mix plus the change between 1940 and 1980 ($\Delta \ln(\frac{S}{U})_{i,1980-1940}$), and substitute into (1'):

$$\Delta Y_{i,2000-1980} = \gamma_{01} + \gamma_{11} \ln(\frac{S}{U})_{i,1940} + \gamma_{12} \Delta \ln(\frac{S}{U})_{i,1980-1940} + \gamma_2 z_{i,1980} + \mu_{i,2000-1980} \tag{1''}$$

Under this additional assumption, $z$ will not bias the coefficients on each component of $\ln(\frac{S}{U})$ in the same way. (This is also true of reverse causality, which seems more likely create a strong correlation between $z_{i,1980}$ and the part representing *changes* in skill mix leading up to the PC revolution.) We show this in Appendix A. Estimates of (1") can thus falsify that $\ln(\frac{S}{U})_{1980}$ is exogenous: if it is, estimates of $\gamma_{11}$ and $\gamma_{12}$ will be different. This approach is like the conventional IV-equivalent "control function" approach of controlling for the first stage residuals to eliminate the bias in OLS, but generalizes it by not requiring us to take a

[37] 1940 has the advantage of not only long predating the computer revolution - or any electronic computers - but also the massive expansion in college attainment in the U.S. (associated, e.g., with the World War II "G.I." Bill.)

stand on which component is the source of "bad" variation in 1980 skill mix.[38]

A weakness of this approach, as we show in Appendix A, is that it may be low power if the autocorrelation of $z$ and $\ln(\frac{S}{U})$ are similar or if the relationship between them evolves only slowly. A key reason this might be the case if the bias worked through general capital-skill complementarity, which Goldin & Katz (2008) argue has been stable over the entire twentieth century. However, economic historians, including Goldin & Katz, also argue that prior to the twentieth century, capital and skill were better described as substitutes.[39] Taking our cue from this, our second instrument is a *nineteenth* century skill mix measure: the share of 15-19-year-olds enrolled in school in 1880, which we think of as a proxy for the local availability of high schools.[40] High school attainment was at the top of the skill distribution at the turn of the twentieth century, giving access to expanding office jobs (Goldin & Katz, 2008).[41] The fact that this instrument is significantly related to education levels a century later suggests it reflects such education supply differences.[42] The highest values of the instrument include many small midwestern and western cities of the sort where the "high school movement" later is documented to have strongly taken hold (Goldin, 1998),[43] while the lowest values of the instrument are dominated by places near the Mexican border.[44] (We control for Mexican

[38]If $\ln(\frac{S}{U})_{i,1940}$ is a valid instrument with a first-stage coefficient of one, $\Delta\ln(\frac{S}{U})_{i,1980-1940}$ would be the control in the control function. Generalizations of the control function include Garen (1984) who addressed selection bias by controlling for the interaction of predicted residuals with main treatment variable, and Blundell & Powell (2004) who address endogeneity with a binary dependent variable.

[39]Artisans competed with low-skill and capital-intensive factory-produced products. Goldin & Sokolof (1984), motivated by nearly the same model as our paper, show that supplies of *low*-skill labor drove adoption of nineteenth century factory technology.

[40]Constructed using a 10% sample of the 1880 Census of Population (Ruggles et al., 2009). Literacy rates from pre-1940 censuses are sometimes also used as proxies for education (e.g., Leon, 2005). Our 1980 skill measure is indeed positively related to 1880 adult literacy rates, though the relationship with the school enrollment rate is stronger.

[41]Why not use *college* enrollment rates? Moretti (2004) use 19th century university density as an instrument for college share. We used his instrument in previous versions of this paper, but ended up shying away from college-level variation out of a concern that universities have a direct effect on local technology adoption.

[42]Even in 2000 data, the 1880 enrollment rate is significantly associated with rates of high school completion. A possible alternative source of the long correlation is tendency of children's education to be is correlated with their parents', which some evidence suggests is driven by selection (e.g., Black et al., 2005). This seems unlikely to fully account for the observed correlation over a century. The intergenerational education correlation is typically around 0.4 in modern data (confirmed by authors' calculations using National Longitudinal Survey of Youth). The labor force in 1980 is on average at least three generations away from 15-19-year-olds in 1880. Based on this we might expect a correlation smaller than $0.4^3$=0.06; in fact, even with controls, the observed partial is correlation is two to three time larger. This suggests there is some slowly evolving city factor maintaining high levels of education in areas with high 1880 enrollments.

[43]The top 10 cities, from largest to smallest enrollment rates, are: Bellingham, WA; Eugene, OR; Kokomo, IN; Bangor, ME; Lawrence, KS; Rapid City, SD; Boise City, ID; Champaign, IL; Wichita, KS; Iowa City, IA.

[44]The bottom 10 cities, from largest to smallest, are: Lafayette, LA; Laredo, TX; Alexandria, LA;

and immigrant shares directly in the regressions.) Overall, the distribution is fairly uniform, with a standard deviation of about 0.1, and a range of 0 to 0.6.[45]

Our argument for the validity of the 1880 instrument is that whatever forces lead to some areas to provide better high schools in the 1880s (such as civic-mindedness or wealth), it is unlikely to be related in the same way to $z$ as of 1880 as $\ln(\frac{S}{U})_{1980}$ is to $z_{1980}$, since the structure of the economy was so different then. What this approach cannot rule out (and nor could any historical instrument) is that these late nineteenth century education supply differences induced accumulation of $z$ over the twentieth century, and it is $z$ that really caused the changes in $Y$ we observe 1980-2000. More concretely, highly educated cities at the turn of the twentieth century would have been at an advantage with the arrival of modern industrial methods, an advantage they could have (perhaps actively) maintained all the way up to the PC era. Nevertheless, it would be quite remarkable per se to find that noisily measured 1880 school enrollment rates predicted skill mix, PC adoption, and increases in the return to skill more than a century later.

The use of these historical variables comes at cost: they are only available in a subset of metropolitan areas. The 1940 Census identifies particularly few metropolitan areas. To keep the tables simple, we will estimate the relationships in two samples: what we will call the "full" sample – the 217 metropolitan areas which can be identified in the 1880 census – and the "1940" sample – the subset of 151 areas that can also be identified in the 1940 census.[46]

# 4 Empirical results

Our exploration of the relationships implied by the theory will broadly follow the order of the propositions and corollaries in the model section. We begin by examining the determinants of technology adoption as measured by the diffusion of PCs (Proposition 1). The first results echo those presented in Doms & Lewis (2006) by looking at the link between PC adoption and the local supply of skill. We then go further by examining whether the return

Brownsville, TX; Titusville, FL; McAllen, TX; Las Cruces, NM; El Paso, TX; Amarillo, TX; Wichita Falls, TX. Tuscon and Phoenix are also near the bottom. In contrast, coastal California cities are in the middle of the distribution, while inland ones tend to have above average enrollment rates.

[45]Census figures from this era are known to vastly overstate full-time school enrollment because any (even trivial) period of enrollment is counted as "enrolled."

[46]The 1880 Census identifies county, but not all of the counties that make up modern metropolitan areas existed by 1880. The 1940 census identifies "state economic areas" which are aggregates of several counties. As a result, many smaller metropolitan areas cannot be identified in the 1940 census.

to skill in 1980, that is at the beginning of the PC diffusion process, is negatively related to the intensity of PC use in 2000 (Corollary 1). We next turn to examining implications of the model for changes in the return to skill, as measured by the ratio of college wages to high school wages. In particular, we explore whether the return to skill increased most where skill was initially most abundant (Proposition 2). The model also implies that the initially negative relationship between the return to and the supply of skill will become less negative and possibly entirely disappear. However, as indicated by Proposition 3, this relationship is predicted to remain non-positive. We then go back and examine whether the return to skill increased most where PCs were adopted most aggressively (Corollary 2), while simultaneously verifying whether the relationship between city-level PC intensity and the return to skill exhibits a weakly negative relationship (Corollary 3). Recall that these propositions were shown to be robust to allowing local supplies of skill to react to changes in the returns to skill (Proposition 5).

## 4.1 PC Adoption and Local Market Conditions

Table 1 reports regression results motivated by Proposition 1 and Table 2 reports results motivated by Corollary 1. Table 1 examines whether PCs were adopted more intensively in cities where the initial skill supply was high. The dependent variable is the average ratio of PCs per employee within detailed industry by plant size cells as of 2000, as described in the data section. This variable is treated as $\Delta Y_i$ in Equation (1) since PC intensity in 1980 was near zero.

Columns (1) and (2) of Table 1 report the slope coefficient from a regression on $\ln(\frac{S}{U})$ – the log ratio of college equivalents to non-college equivalents – as of 1980 in the full and 1940 samples, respectively, without additional controls. There is a strong positive correlation that is similar in magnitude in both samples – a 10% increase in skills is associated with roughly 1.5 extra PCs per 100 population. As a first attempt to explore the endogeneity of $\ln(\frac{S}{U})_{i,1980}$ we add controls in columns (3) and (4). This lowers the point estimate a bit, though the confidence intervals overlap.

In column (5) we exploit the fact that $\ln(\frac{S}{U})_{i,1980}$ can be expressed at the sum of $\ln(\frac{S}{U})_{i,1940}$ and $\Delta \ln(\frac{S}{U})_{i,1980-1940}$ to attempt to falsify the exogeneity of $\ln(\frac{S}{U})_{i,1980}$. The estimates on both components of $\ln(\frac{S}{U})_{i,1980}$ are similar, and are not statistically distinguished from one another (p-value = .15). This first of all suggests that "reverse causality" is unlikely to be driving the observed correlation, since we would expect this to be much more severe for $\Delta \ln(\frac{S}{U})_{i,1980-1940}$ than for $\ln(\frac{S}{U})_{i,1940}$. Workers likely had some sense that the arrival of PCs would raise demand for skills, but they seem not to have anticipated its differential impact on demand for skills across different labor markets. More broadly, if the estimates are driven by unobserved third factors, only if these unobserved factors have a stable relationship with $\ln(\frac{S}{U})$ from 1940 to 1980 would expect these coefficients to be similar.[47]

Some argue that the degree of capital-skill complementarity has been stable over the entire twentieth century (Goldin & Katz, 2008), which provides a reason the test in column (5) could be low-powered. (See Appendix A.) To address this, we turn to our other instrument which comes from before the era of capital-skill complementarity: 1880 enrollment rates of 15-19-year-olds. This variable is surprisingly correlated with 1980 skill mix: the first stage coefficient on 1880 enrollment is 0.45 – a ten percentage point higher enrollment rate in 1880 (roughly one standard deviation) is associated with 4.5% greater skills in 1980 – with an F-stat of 15.75. It is also correlated with PC adoption from 1980-2000, with the IV coefficient not statistically distinguished from the OLS estimate (Hausman F-stat on $\ln(\frac{S}{U}) = 0.9$).[48] This does not rule out that other forces generate this 120 year correlation though, recall, a simple industry mix story is insufficient, as our PC intensity measure is adjusted for detailed industry.

In Table 2, we turn to examining the link between the adoption of PCs over the period 1980-2000 and the initial college-high school wage gap, the "price of" or "return to" skill. Corollary 2 implies that PCs should be adopted more intensively in cities where the price of skill is initially low, since such cities have a comparative advantage in adopting the new skill-intensive technology. In column (1) and (2) we show the relationship between PC intensity in 2000 and 1980 skill prices in the full and 1940 samples, respectively, without additional controls. There is indeed a significant negative relationship between the initial price of skill and the adoption of PCs. Columns (3) and (4) add industry and city controls to the 1940 sample. This reduces the coefficient, though not in a statistical sense (also true in the full sample, shown in column 7).

We do not believe that the differences in the return to skill in 1980 are exogenous with respect

[47] An additional requirement for the coefficients to be the same is that these unobserved factors and $\ln(\frac{S}{U})$ have a similar autocorrelation 1940-1980: see Appendix A.

[48] For all Hausman tests we used OLS standard errors, rather than heteroskedasticity-consistent standard errors.

to unobserved factors that may favor PC adoption, so the results are only suggestive.[49] Accordingly, the remaining columns consider different IV estimates, taking advantage again of the idea that 1980 *skill* differences may be exogenous (which we failed to reject in Table 1). In Column (5), we use $\ln(\frac{S}{U})_{i,1980}$ as the instrument. As we will see below (Table 4), this is strongly negatively related to the return to college in 1980 - providing more support for the idea that 1980 skill mix differences are supply driven - and the first-stage F-stat is 50.11. The IV estimate is larger (significantly so) than the OLS estimate. In Column (6) we break up the variation in $\ln(\frac{S}{U})_{i,1980}$ into $\Delta \ln(\frac{S}{U})_{i,1980-1940}$ and $\ln(\frac{S}{U})_{i,1940}$. The IV estimate is nearly the same, and the data fail to reject the overidentifying variation is correlated with the error (p-value=0.15). Again, this might be a low power test if the sources of bias derive from something related to capital-skill complementarity from before the PC era, which may have been stable to as far back as the early twentieth century. However, our second, nineteenth century instrument gives similar results (column 8).

Why are the IV estimates larger? We think a plausible explanation is that some cities are more innovative in the sense they possess some unobserved factor which raises both the return to skill in 1980 and subsequent PC adoption. Such a factor would bias toward zero what would otherwise be a negative relationship between PC adoption and the initial price of skill. The fact that the Table 1 estimates are not sensitive to the use of our instruments, while the estimates in Table 2 are, suggests that this unobserved favor is not very strongly correlated with historical education differences.[50]

## 4.2 Changes in the Return to Skill and Initial Skill Level

A central prediction of the model presented in Section 2 is that, during a technological transition, the return to skill should increase the most where supply of skill is initially highest (Proposition 2). Furthermore, such a relationship should arise due to a flattening of the relationship between the return to skill and the supply of skill (Figure 2). Accordingly, in Table 3a, we report results from regressions aimed at documenting the relationship between the change in the return to skill over the period 1980-2000 and the level of skill in 1980.

[49] One might legitimately ask what maintains differences in skill prices across markets; Black et al. (2005) argue that the differential valuation of certain amenities may allow this. This suggests the possibility that some amenities might be valid instruments for returns, an approach we considered.

[50] Unobserved factors positively correlated with PC adoption and historical education differences could, alternatively, be biasing up IV estimates in Table 2, but they would then do the same to estimates in Table 1, where we failed to reject the difference between OLS and IV.

In columns (1) and (2) we report OLS estimates without any additional controls in the full and 1940 samples. These estimates suggest the return rose significantly faster in initially more educated markets: a 10% increase in skills in 1980 (one-third of a standard deviation) is associated with 0.7 percentage point larger change in the return to college 1980-2000. A scatter plot of this relationship is shown in Figure 5.[51]

The effect is slightly larger when the city and industry controls are added to the 1940 sample in column (3), though not in a statistical sense. Column (4) and the remaining columns also correct for the endogenous selection into particular metro areas using Dahl's (2002) flexible control function approach. The idea behind Dahl's approach is that place of birth is a strong predictor of a person's eventual labor market, but within demographic categories, place of birth is orthogonal to the potential return to skill (the maintained assumption). In practice, the correction involves controlling for a polynomial in the probability that an individual is observed in a particular metropolitan area, estimated from the observed distribution of individuals across metro areas by demographic category (age x race x gender) and state or country of birth.[52] Consistent with our argument that endogeneity is not a major problem here, the selection correction has a negligible effect on the results.[53]

Column (5) decomposes initial skill mix into the 1940 level and the 1940-1980 change. As before, we expect that any endogeneity of skill mix should manifest itself in the form of different coefficients on the two components. The estimated effects of both components are similar, and we fail to reject that they are the same. To try to rule out that this is because of stability in the source of bias going back to 1940, we also include IV estimates using our 1880 instrument. This is shown in column (7), with comparable OLS estimates in column (6). We fail to reject that OLS and IV are the same (Hausman F = 1.94) though the IV estimate is larger.

In the last two columns we examine what has happened since 2000 by making the end year 2008 (constructed using data from stacked 2006-2008 American Community Surveys from Ruggles et al., 2009). Some theories say that the increase in the skill premium is temporary to an "implementation phase" of the the new technology (Galor & Moav, 2000; Greenwood & Yorukoglu, 1997). If so, our data suggest that the implementation phase is not over by

[51] San Francisco appears to be an outlier in this figure, and although it does contribute positively to the slope, it does not by itself drive the results.

[52] The details of how this is implemented mirror Beaudry, Green and Sand (2007) and are described in Appendix B.

[53] The control function is nevertheless significantly related to returns.

2008: the skill-return relationship continued to rise modestly through 2008. We do not have the data to show it, but we suspect PC intensity has also continued to rise.

Another issue is whether the relationship in Table 3a simply reflects trends in the return to skill which predates the PC revolution. We might be worried, for example, that earlier generations of technologies favoring skilled workers were adopted more quickly in these permanently more "innovative" cities.[54] We investigate this empirically in Table 3b. Obtaining reliable estimates of the return to college by area prior to 1980 turns out to be quite difficult because wage samples for years before 1980 tend to be small. To deal with this, we have combined the 1940 and 1950 censuses, which report geography in identical detail. Even with the stacked data, the number of wage observations, especially for college graduates, is quite small in many metropolitan areas.[55] To increase the sample further, rather than including only those with exactly 12 and exactly 16 years of education in the construction of return to skill (as we do in all other tables), we instead estimated the linear return to college equivalent status (equal to one for four-year grads, 0.5 for some college but no degree, and zero otherwise) using a sample of all individuals with at least 11 years of education. To deal with the skill (education) heterogeneity among people with the same college equivalent status, we conditioned our city-specific estimates on a common (i.e., not city-specific) linear control for years of education. We also conditioned on a quartic in potential work experience, a dummy for female and a dummy for foreign-born.

Panel A of Table 3b reports the results of regressions of change in the adjusted return in between 1940/50 and 1980 on 1980 skills. To make sure our earlier results are robust to the change in adjustment methods, Panel B Table 3b report regressions of the change in the return between 1980-2000, adjusted with the same methods, on 1980 skills. In Panel B, a significant, positive relationship remains with and without controls, and the point estimates are surprisingly similar to the estimates using the original methods. In contrast, there is no evidence of a positive relationship between 1980 skill share and the change in returns between 1940/50 and 1980. Thus, it appears that the relationship we are uncovering is robust to methodological changes, and is exclusive to the recent period of rapid technological

[54]Related to this, Goldin & Katz (2008) argue that computers are just the last example of a century-long trend of improving machinery raising demand for skills. If there is, say, learning-by-doing associated with older technologies which complements newer ones, then a locality's prior technology adoption would be an omitted variable correlated with skill. To attempt to refute this sort of view, Doms & Lewis (2006) show that conditional on education levels, the computer-intensity of an area in the late 1970s does not predict PC-intensity later.

[55]In the 1940 and 1950 public-use Census samples, only a small subset of so-called "sample-line" respondents were asked to report their wages.

change.[56] This supports the earlier evidence that skill mix in 1980 may have been exogenous to other forces affecting the return to skill. Remarkably, column (4) shows that even 1880 school enrollments predict increases in returns from 1980 to 2000 but not before.[57]

While the patterns highlighted in Tables 3a and 3b are consistent with our model, they could be consistent for the wrong reason. In particular, greater increases in the return to skill arising where skill is most abundant could arise from three different underlying movements in the return-supply relationship: from a positive relationship becoming more positive, from a negative relationship turning positive, or from a negative relationship becoming less negative but not turning positive. Only the third case is consistent with our model of technological revolutions (Proposition 4).

To examine these implications, Table 4 reports estimates of the relationship between return to skill and the skill mix in the years 1940/50, 1980, 2000, and 2008.[58] With no controls, in columns (1) - (4), the estimated relationship between the return to skill and our skill measure goes from from -0.05 in 1940 to -0.06 in 1980 to a statistically insignificant relationships in 2000 and 2008. Adding controls, in Columns (5)-(8), and this "flattening" pattern is strengthened. In short, there appears to be a relatively stable negative relationship between the return to skill and skill supply that disappears sometime after 1980.[59] We have also tried using 1940 skill mix or 1880 enrollment rates as instruments: this makes the standard errors and point estimates bigger, but shows the same pattern of flattening to zero relationship. So the data do appear to be following closely the pattern predicted by the model. Note also that OLS produces significant negative estimates in the pre-period, consistent with at least part of the skill mix variation being supply-driven.[60]

[56] We also tried looking at the 1970 Census, where the sample is a bit larger but the number of identifiable metropolitan areas is smaller. Using these data, we also find no significant association between pre-1980 change in the return and 1980 skills. See note below. (Metro areas are not identifiable in the 1960 Census.)

[57] The first stage is on the weak side in the smaller 1940 sample – 8.07 – which means it is more appropriate to rely on the reduced form p-value for statistical significance. It is 0.079 for the 1980-2000 outcome.

[58] The 1940/50 returns are estimated with the alternative adjustment methods of Table 3b, while the other are not. It turns out, however, the estimates in other years are robust to the alternative methodology.

[59] We also tried estimating this for 1970. With controls, and using the 132 of our metropolitan areas that can be constructed with 1970 census geographic variables, the relationship between the return and contemporaneous skill ratios has a slope (standard error) of -0.161 (0.0299) in 1970 and -0.177 (0.0248) in 1980.

[60] The point estimates for 1980 correspond to an elasticity of substitution between college and high school labor in the range of 6-16 (inverse of point estimate). This is smaller than estimates of the elasticity obtained from aggregate data (e.g., Katz & Murphy, 1992) but similar to estimates based on cross-area variation, such as from research on the local labor market impact of immigration (e.g., Card, 2001). The IV estimates (not shown in table) imply a smaller elasticity on the order of 3-4, closer to the aggregate estimates.

The data are also consistent with Corollary 2 and 3. However, for brevity, we do not provide a set of tables for these results (available from the authors upon request). Not too surprisingly, we do find that cities which adopted PC more intensively saw the return to skill increase most, consistent with Corollary 2. The more surprising result relates to the link between the return to skill in 2000 and the use of PCs in 2000. The theory suggests that we should *not* see a positive link between these two, that is, cities that use PCs intensively should not, on average, have a higher return to skill. Consistent with this, a regression of the return to skill in 2000 on PCs per worker in 2000 produces a slope coefficient of almost zero (0.001 with a heteroskedasticity-robust standard error of 0.07). Controls and instruments reduce somewhat the precision of this estimate but do not make the relationship positive. Note how this contrasts with individual-level data in which the return is significantly higher for more educated individuals who use computers (Krueger, 1993; Valletta, 2006).[61] This observation thus provides considerable support to the notion that the adoption of PCs should be seen as endogenous process that responds to the low cost of skill.

## 4.3 What Do Our Data Say About Related Models?

In section 2.2 we described several models related to ours. What do the data say about these models?

A major focus of models of computers as a new "General Purpose Technology" is the fact that there is an initial phase in which the arrival of the new technology leads to decreased productivity, as resources are diverted to complementary R&D investments. To see if our data are consistent with this, we constructed a measure of *average* wages adjusted using similar methods as we used for the return to skill.[62] These data provide some evidence

[61] Valletta finds that Krueger's result is sensitive to specification. However, in both 2001 and 2003 Current Population Survey (CPS) computer-use supplements, Valletta also finds the return to college is larger for computer users. We have confirmed, using our approach, that the difference is not to due to the data source but to the level of aggregation. In particular, in the subsample of 163 of our metro areas that have data on computer use and wages in the 2001 CPS supplement, the adjusted (using the same controls – See Appendix B) college-high school ln wage gap is 15 percentage points larger for computer users than non-users (area-clustered standard error = 0.042). Nevertheless, when this same CPS computer use measure is aggregated to the metro area level, it has zero relationship with the return to college: the coefficient (heteroskedasticity-robust standard error) on our estimate of the 2000 return is -0.026 (0.032).

[62] In particular, in the sample of workers with at least 11 years of education, we regressed the real ln hourly wage on a quartic in potential experience, linear years of education in addition to education category dummies – specifically, dummies for exactly high school graduate, some college but no degree, exactly four-year college graduate, and postgraduate (dropouts are the excluded group) – dummies for white, female, married, age (four categories), born after 1950, born after 1950 interacted with completed at least four

that places which adopted PCs in the 1990s had slower productivity growth in the 1980s. Conditional on our city and industry controls, places which were more educated (by our measure) in 1980 (or using the 1880, though not significantly, or 1940 instrument) had significantly slower wage growth in the 1980s and significantly faster wage growth in the 1990s, consistent with a productivity slowdown in advance of their faster PC adoption.[63] However, unlike our other results in this paper, this is sensitive to controls. Without the controls, 1980 skill share is associated with faster wage growth in both decades.

What does our evidence say about the "skill-biased technical change" (SBTC) view of computers, namely that the falling price of skill-complementary computers has raised skilled relative wages? If one defines capital-skill complementarity generally enough, it can encompass as a special case the model of endogenous adoption we have presented, in which case the two are not distinguishable. More narrowly, however, the models of capital-skill complementarity typically used in SBTC studies appear not to be able to fully account for the cross-city facts uncovered in this paper.[64]

The standard model in SBTC papers are versions (sometimes simplified, sometimes nested slightly differently) of the nested CES model presented in Krusell et al. (2000):

$$g(PC, S, U) = (c(PC^{\nu} + S^{\nu})^{\frac{\mu}{\nu}} + (1-c)U^{\mu})^{\frac{1}{\mu}}$$

In this framework, the driving force is a fall in the price of computing equipment, $r^{PC}$. To bring this to the local market, assume, as we do in this paper, that PCs are supplied elastically at $r^{PC}$; also assume $\mu > \nu$ (PC-skill complementarity) so that $\frac{w_i^S}{w_i^U} = \frac{g_S(PC_i,S_i,U_i)}{g_U(PC_i,S_i,U_i)}$ is decreasing in $r^{PC}$ . In this model, like in ours, an increase in the relative supply of skilled labor will also increase adoption of PCs. The distinction is what happens to relative wages. In general, this functional form cannot replicate our finding that the arrival of PCs is associated with reduced sensitivity of relative wages to relative supply. Note that this not the

years of college, and metropolitan area. The coefficients on the metropolitan area dummies are the adjusted average wages we used.

[63] The estimated effect is fairly small – a one standard deviation increase in skills (0.3 log points) is associated with about 1.3 percentage points slower wage growth (one-sixth of a standard deviation) in the 1980s, and a similar magnitude faster growth in the 1990s.

[64] A possible reason for this is that cross-area analysis has not been a major part of the SBTC literature, which has focused on explaining aggregate changes in the wage distribution. An exception is Autor & Dorn (2009), who apply a generalization of Autor, Levy, and Murnane (2003) to US labor markets. In their view, initial differences in occupation mix will drive local changes in the wage distribution.

same as requiring $\frac{\partial \frac{w_i^S}{w_i^U}(r^{PC}, \frac{S}{U})}{\partial r^{PC}} < 0$, but rather that the *second* derivative $\frac{\partial^2 \frac{w_i^S}{w_i^U}(r^{PC}, \frac{S}{U})}{\partial r^{PC} \partial \frac{S}{U}} < 0$.[65]

There does not appear to be an easy way to sign this second derivative. However, in the special Cobb-Douglass case (in the limit approaching $\mu = 0$), one can show that the second derivative is actually *positive.* In the more general case where $\mu \neq 1$, we examined the sign of $\frac{\partial^2 \frac{w_i^S}{w_i^U}(r^{PC}, \frac{S}{U})}{\partial r^{PC} \partial \frac{S}{U}}$ numerically. Looking over a very wide range of parameters considered relevant in the literature, we found no cases where this second derivative was negative.[66]

Although we recognize that this is not a proof, it appears that embedding PC capital in a nested CES production function and having its price fall is not sufficient to describe the patterns of changes in the return to skill across markets 1980-2000. Some additional elements appear necessary. For example, to explain the patterns reported in Tables 3b and 4 in a one production function framework, it is likely to be necessary to drop the assumption of constant elasticities of substitution over time. Note that one of the properties of the endogenous technology adoption model is to offer a structured explanation to why certain elasticities will depart from the conventional norms precise a time of great technological revolutions.

## 4.4 Quantifying Effects

Our model of endogenous technology adoption has the property that localities with different levels of educational attainment should adjust differently to a technology revolution, with more educated cities adopting the new technology more aggressively and thereby experiencing a greater increase in the return to skill. The evidence present in the tables provides support for this view. However, from these tables it is difficult to assess the importance of this process in relation to the overall increase in the return to education observed over the period 1980-2000.

It is interesting to ask, for example, what our estimates imply regarding the magnitude of

[65] Since the return to skill is determined by properties of the first order derivative of the production structure, $\frac{\partial^2 \frac{w_i^S}{w_i^U}(r^{PC}, \frac{S}{U})}{\partial r^{PC} \partial \frac{S}{U}}$ implicitly depends on the third order derivative of the production structure.

[66] In this discussion, we focused on the sign of $\frac{\partial^2 \frac{w_i^S}{w_i^U}(r^{PC}, \frac{S}{U})}{\partial r^{PC} \partial \frac{S}{U}}$, while in the empirical section we focused on the effects on changes in $\ln \frac{w_i^S}{w_i^U}$. To be consistent, we have verified that the observation of greater increases in skill where skill is most abundant is robust to whether or not skill is measured as $\frac{w_i^S}{w_i^U}$ or $\ln \frac{w_i^S}{w_i^U}$.

increased return to education in a highly educated city versus a poorly educated city. To take two extremes, we can compare the predicted outcomes for Hickory, NC, one of our sample's least educated areas in 1980, and Tallahassee FL, one of our sample's most educated areas. (See Figure 4.) In 1980, the fraction of college workers was 16% in Hickory, while it was 44% in Tallahassee.[67] Taking the modal coefficient estimate of about 0.1 from Tables 3a and 3b, our estimates imply that the return to college would have increased 14 more log points in a city like Tallahassee than in one like Hickory. This gap is large – it is similar in magnitude to the increase in the (unadjusted) return in the average city in our sample between 1980 and 2000.

More broadly, a one standard deviation increase in 1980 skills (that is $\ln(\frac{S}{U})$), or about 30 log points, is estimated to lead to three percentage points faster growth in the return to college over 1980-2000, or about half a standard deviation increase in the return. A 30 log point increase is quite reasonable in the time series as well – the average area raised its college share by 65 log points between 1980 and 2000.

# 5 Alternative Explanation: Integration?

Our model's predictions regarding changes in factor prices are identical to those of a model of a small economy going from autarky to free trade. Although there are no tariff barriers between U.S. markets, trade between US cities is not costless. So a sudden fall in the costs of trading between U.S. labor markets over the period 1980-2000 - perhaps enabled by information technology - could be the reason for some of the data patterns we have highlighted. In this section we explore whether observed changes in industrial composition provides support for this alternative explanation.

If a fall in inter-city trade cost is an important driving force behind wage movements observed over the period 1980 to 2000, we should observe industries reallocating themselves across markets to better align with comparative advantage. In other words, skill-intensive industries should move to areas where skill is abundant and the price of skill is low. A useful

[67] That Tallahassee was the most educated city in 1980 may be unexpected to some readers, but recall that our measure of college share includes half of individuals with some college education but no four-year degree. Tallahassee also has a lot of university capacity for an area of its size, including Florida State. A handful of the most educated markets (e.g., Columbia, MO; Madison, WI) might also be considered "university towns." It is interesting, however, that returns do not appear to grow any faster in university towns than in other similarly-educated markets.

empirical framework for evaluating this is provided by Card & Lewis's (2005) decomposition of cross-sectional differences in skill mix into "between" and "within" industry components. This decomposition is applied to a city's "excess college share" defined as the difference in college share in city $c$, $\frac{S_c}{S_c+U_c}$, and the college share for the country as a whole, $\frac{S}{S+U}$.[68] The decomposition takes the following form:

$$\frac{S_c}{S_c+U_c}-\frac{S}{S+U}=\sum_j \frac{S_j}{S_j+U_j}\left(\frac{S_{c,j}+U_{c,j}}{S_c+U_c}-\frac{S_j+U_j}{S+U}\right)+\sum_j \frac{S_{c,j}+U_{c,j}}{S_c+U_c}\left(\frac{S_{c,j}}{S_{c,j}+U_{c,j}}-\frac{S_j}{S_j+U_j}\right)$$

In the above decomposition, industries are indexed by $j$, so $\frac{S_j}{S_j+U_j}$ represents the college-intensity of sector $j$ in the nation as whole, and $\frac{S_{c,j}}{S_{c,j}+U_{c,j}}$ is the college-intensity of sector $j$ in city $c$. A city's excess college share is decomposed above into, first, a "between" term which reflects the city's excess of college-intensive industries and, second, a "within" term which reflects the city's excess of college-intensity within sectors.

Previous research has found that in cross-section, very little of the differences in skill mix across US cities are accounted for by the between industry term, i.e., differences in industry mix (e.g. Lewis, 2003). Here, we are interested in whether *changes* in industry mix reflect increased integration between 1980 and 2000. The change in the industry mix portion of the between industry term, given below, provides a measure of the extent to which changes in a city's industrial structure reflects a concentration of jobs in high skill industries.

$$\sum_j \frac{S_j}{S_j+U_j}\left[\Delta\left(\frac{S_{c,j}+U_{c,j}}{S_c+U_c}\right)-\Delta\left(\frac{S_j+U_j}{S+U}\right)\right] \qquad (2)$$

For the purpose of empirical implementation, $\frac{S_j}{S_j+U_j}$ is calculated as the college equivalent share for industry $j$ in 2000 and the term in brackets is equal to the change in employment share in industry $j$ in city $c$ compared to industry $j$ in the nation as a whole between 1980 and 2000. If trade integration is occurring rapidly between 1980 and 2000, then this statistic should be high for cities with an initially high college share, as college-intensive industries should disproportionately relocate to such cities to take advantage of their abundance (low price) of skilled labor, and it will be low for cities with initially low college share. In the extreme where trade integration leads to shifts in industry mix which fully absorb initial

[68] Cities were previously indexed with $i$, but are indexed with $c$ in this section to avoid confusion of city and industry indexes.

differences in skill mix (and nothing else changes), this statistic will be exactly equal to each city's initial excess college share $\frac{S_c}{S_c+U_c} - \frac{S}{S+U}$.

In order to evaluate the extent to which this is the case, Figure 6 plots this between-industry statistic (given in equation 2) against 1980 excess college share and fits the relationship between the two. Differences in the between-industry statistic are first of all small compared to the differences in college share – note the difference in scale on the y- and x-axis. There is also little sign that the growth in a city's college-intensive industries has a positive relationship with its initial college share. One of the least educated cities in our sample, Scranton, PA, has one of the larger increases in college-intensive industries between 1980 and 2000; one of the most educated, Tallahasee, has one of the smallest increases. More generally, the coefficient from the regression line, which can be interpreted as the proportion of initial differences in college share absorbed by changes in industry mix, is 0.02 with a standard error of 0.02. Put another way, industry mix is evolving in roughly the manner one would expect if very little additional integration were occurring.[69]

Another test is to examine the relationship between this statistic and each area's initial (1980) return to skill. If trade integration occurs between 1980 and 2000, this statistic should be negatively correlated with the initial return to skill as skill-intensive sectors "move out" (in relative terms) of areas with a high relative price of skill and move to areas with a low relative price of skill. In fact, the relationship between this statistic and the initial return is slightly positive; skill-intensive sectors did not differentially move away from markets where the relative price of skill was high in 1980. Thus, at the level of industry detail available in the 1980 census (roughly three-digit SIC),[70] there is no evidence that cities are "integrating" quickly over this period according to comparative advantage principles based on factor endowments. While some research suggests that integration may show up in changes in product quality below the level of industry detail we can observe (Schott, 2004), it is not very encouraging for the trade integration view that there is virtually no action at the level of industry detail we do observe.

Another difficulty with the increased inter-city trade explanation is that, controlling for industry structure, there should be no systematic differences in PC use across cities. Cities

[69] The regression used the same weights used in the regression presented in tables 1-4. Adding the controls from those tables reduces the coefficient; using the instrument makes the relationship slightly negative, though not statistically significantly so.

[70] Industry codes are more detailed than this in the 2000 Census. We use Lewis (2003)'s mapping between industry codes in the 1980 and 2000 Censuses.

with more educated workers should have more skill- and PC-intensive industries but should not use PCs more intensively within an industry.[71] However, as was documented in Table 1, there is a strong positive link between PC use within industries (our measure of PC use) and the local supply of skill.

What about the integration through movement of labor instead of goods? This is also not a promising explanation. The standard deviation across cities of our skill mix measure hardly changed between 1980 (0.30) and 2000 (0.29); recall from Figure 4 the ordering of cities stayed nearly the same as well.[72] Finally, there is this bottom line: integration in all forms would be expected to decrease the magnitude of cross-city differences in the return to skill. In fact, the cross-city standard deviation in the return to college did not decrease between 1980 and 2000, it rose from 0.050 to 0.057. So while it is probably true that some additional integration occurred between 1980 and 2000, it is unlikely the cause for the observation that the return to skill increased most where skill was initially most abundant.

# 6 Conclusion

The notion that certain periods correspond to technological revolutions is widespread. However, it is often unclear what characteristics should be met for a period to be considered a technological revolution. A common idea is that such a period should reflect a paradigm shift in the method of production. Accordingly, in this paper we used a simple neo-classical model of adjustment to a new technological paradigm to highlight a set of implications that could define a technological revolution. We then examined whether these implications were observed in the US over the period 1980-2000 when the PC was introduced and diffused. One of the main implications of our model is that the link between supply of factors and their returns should be quite different during a technological revolution than in the long run. As we know from many studies, over the long run the supply of skill places downward pressure on the return to skill. However, as implied by the model, during a period of adjustment to a new technological paradigm that is skill-biased this should not generally be the case. Instead, during such a period, we should observe that the return to skill increases most where skill is most abundant, and we should see increases in skill lead to a faster diffusion of the new

[71] In the case of the PC data, industry is observed at an even higher level of detail than in the census - four-digit SIC. Even controlling for this, there are significant differences in PC use across areas that are correlated with college share.

[72] As noted earlier, by other measures, skill mix differences across cities diverged between 1980 and 2000.

technology without placing downward pressure on the return. Moreover, the model predicts that endogenous diffusion should not lead to a situation where the return to skill is highest in localities that have adopted the new technology most aggressively. Using a combination of data sources, we found evidence supporting all of these predictions. For this reason, the period 1980-2000 may well deserve the designation of a technological revolution.

Is it important to know whether a period is a technological revolution? We believe that the answer is "yes," since such knowledge is useful to better guide and evaluate policies. For example, the arrival of a major skill-biased technology as modeled here will cause the return to skill to increase and therefore lead to increased wage inequality. From a historical perspective, a reasonable policy to combat such increased inequality is to favor greater accumulation of skill. However, during a technological revolution, as we have shown, an increase in the supply of skill is unlikely to reduce wage inequality as the increase in skill instead acts simply to accelerate the adoption of the skill-biased technology. In the absence of any recognition that the period may be a technological revolution, and that therefore historical relations may be temporarily inactive, one may evaluate policies inappropriately and make wrong choices. In particular, during a period of a technological revolution, our analysis suggests that policy makers need to be very patient if they want to see increases in supply have negative impacts on prices, and may need to choose alternative policies if they desire immediate results.

## References


Autor, David, and David Dorn (2009). "Inequality and Specialization: The Growth of Low-Skill Service Jobs in the United States," NBER Working Paper #15150, July.

Autor, David, Lawrence F. Katz and Melissa S. Kearney (2006). "The Polarization of the U.S. Labor Market." *American Economic Review* 96(2): May, pp. 189-94.

Autor, David H., Frank Levy and Richard J. Murnane (2003). "The Skill Content of Recent Technological Change: An Empirical Exploration," *Quarterly Journal of Economics*, 118 (4): November 2003, pp. 1279-1334.

Bartel, Ann P and Frank R. Lichtenberg (1987). "The Comparative Advantage of Educated Workers in Implementing New Technology." *Review of Economics and Statistics* 69 (1): February, pp. 1-11.

Basu, Susanto and David Weil (1998). "Appropriate Technology and Growth," *Quarterly Journal of Economics* 113(4): November, pp. 1025-54.

Basu, Susanto and John Fernald (2008). "Information and Communications Technology as a General Purpose Technology: Evidence from US Industry Data." *Economic Review*. San Francisco: Federal Reserve Bank of San Francisco.

Black, Sandra, Paul J. Devereux, Kjell G. Salvanes (2005). "Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital." *The American Economic Review* 95(1): March, pp. 437-449.

Blundell, Richard W. and James L. Powell (2004). "Endogeneity in Semiparametric Binary Response Models." *Review of Economic Studies* 71(3): July, pp. 655-79.

Beaudry, Paul and David Green (2003). "Wages and Employment in the United States and Germany: What Explains the Differences?" *American Economic Review* 93(3): June, pp. 573-602.

Beaudry, Paul and David Green (2005). "Changes in U.S. Wages, 1976-2000: Ongoing Skill Bias or Major Technological Change?" *Journal of Labor Economics* 23(3): July, pp. 609-648.

Beaudry, Paul, David Green, and Benjamin Sand (2007). "Spill-Overs from Good Jobs," NBER Working Paper #13006, April.

Berry, Chirstopher R. and Edward L. Glaeser (2005). "The Divergence of Human Capital Levels Across Cities." National Bureau of Economic Research Working Paper #11617, September.

Black, Dan, N. Kolesnikova and L. Taylor (2004). "Understanding the Returns to Education when Wages and Prices Vary by Location," mimeo, Carnegie- Mellon University.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes (2005). "Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital." *American Economic Review* 95 (1): March, pp. 437-49.

Bresnahan, T., and M. Trajtenberg (1995). "General Purpose Technologies: Engines of Growth?" *Journal of Econometrics* 65(1): pp. 83-108.

Card, David (2001). "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration." *Journal of Labor Economics* 19(1): January, pp. 22-64.

Card, David and John E. DiNardo (2002). "Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles." *Journal of Labor Economics*, 20 (4): October 2002, pp. 733-83.

Card, David and Ethan Lewis (2005). "The Diffusion of Mexican Immigrants During the 1990s: Explanations and Impacts." NBER Working Paper #11552, August.

Caselli, Francesco (1999). "Technological Revolutions." *American Economic Review* 89(1): May, pp. 78-102.

Dahl, Gordon (2002). "Mobility and the Return to Education: Testing a Roy Model with Multiple Markets." *Econometrica* 70(6): pp. 2367-2420.

Doms, Mark and Ethan Lewis (2006). "The Diffusion of Personal Computers Across the U.S., 1990-2002." Working Paper Series 2006-18, Federal Reserve Bank of San Francisco.

Galor, Oded and Omer Moav (2000). "Ability-Biased Technological Transition, Wage Inequality, and Economic Growth." *The Quarterly Journal of Economics* 115(2): May, pp. 469-497.

Garen, John (1984). "Returns to Schooling: A Selectivity Bias Approach with a Continuous Choice Variable." *Econometrica* 54(2): September, pp. 1199-1218.

Goldin, Claudia (1998). "America's Graduation from High School: The Evolution and Spread of Secondary Schooling in the Twentieth Century." *The Journal of Economic History* 58 (2): June, pp. 345-374.

Goldin, Claudia and Larry Katz (2008). *The Race Between Education and Technology.* Cambridge and London: Harvard University Press, Belknap Press.

Goldin, Claudia and Kenneth Sokoloff (1984). "The Relative Productivity Hypothesis of Industrialization: The American Case, 1820 to 1850." *Quarterly Journal of Economics* 99(3): August, pp. 461-488.

Greenwood, Jeremy and Mehmet Yorukoglu (1997). "1974." *Carneie-Rochester Conference Series on Public Policy* 46: pp. 49-95.

Habakkuk, H. J. (1962). *American and British Technology in the Nineteenth Century.* Cambridge: Cambridge University Press.

Helpman, Elhanan and Manuel Trajtenberg (1998). "A Time to Sow and a Time to Reap: Growth Based on General Purpose Technologies" in Helpman, Elhanan, ed., *General Purpose Technologies and Economic Growth.* Cambridge, MA: MIT Press.

Katz, Lawrence F. and Kevin M. Murphy (1992). "Changes in Relative Wages, 1963-1987: Supply and Demand Factors." *Quarterly Journal of Economics* 107(1): February, pp. 291-303.

Krueger, Alan (1993). "How Computers Have Changed the Wage Structure: Evidence from Microdata 1984-1989." *Quarterly Journal of Economics* 108(1): pp. 33-60.

Krusell,Per, Lee E. Ohanian, Jose - Victor Rios-Rull, and Giovanni L. Violante (2000). "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis." *Econometrica* 68(4): September, pp. 1029-1055.

Leon, Alex (2005). "Does 'Ethnic Capital' Matter? Identifying Peer Effects in the Intergenerational Transmission of Ethnic Differentials." Mimeo, University of Pittsburgh, August.

Lewis, Ethan (2003). "Local, Open Economies Within the US: How Do Industries Respond to Immigration?" Federal Reserve Bank of Philadelphia Working Paper #04-01, December.

Lewis, Ethan (2005). "Immigration, Skill Mix, and the Choice of Technique." Federal Reserve Bank of Philadelphia Working Paper #05-08, May.

Lipsey, Richard, Cliff Bekar, and Kenneth Carlaw (1998). "What Requires Explanation?" in Helpman, Elhanan, ed., *General Purpose Technologies and Economic Growth.* Cambridge, MA: MIT Press.

Moretti, Enrico (2004). "Estimating the Social Return to Higher Education: Evidence From Longitudinal and Repeated Cross-Sectional Data." *Journal of Econometrics* 121(1-2): July-August, pp. 175-212.

Nelson, Richard R. and Edmund S. Phelps (1966). "Investment in Humans, Technological Diffusion, and Economic Growth." *American Economic Review* 56(2): March, pp. 69-75.

Saint-Paul, Giles (2008). *Innovation and Inequality: How Does Technical Progress Affect Workers?*. Princeton and Oxford: Princeton University Press.

Schott, Peter K. (2004). "Across-Product versus Within-Product Specialization in International Trade." *Quarterly Journal of Economics* 119(2): May, pp. 647-678.

Valletta, Robert (2006). "Computer Use and the U.S. Wage Distribution, 1984-2003." Federal Reserve Bank of San Francisco Working Paper #2006-34, October.

Wozniak, Abigail (2006). "Why Are College Graduates More Responsive to Distant Labor Market Opportunities?" Mimeo, University of Notre Dame.

Zeira, Joseph (1998). "Workers, Machines, and Economic Growth." *Quarterly Journal of Economics* 113(4): pp. 1091-1117.

Table 1: PCs per Worker, 2000, and 1980 Education

| Estimation: | OLS | OLS | OLS | OLS | OLS | OLS | IV[1] |
|---|---|---|---|---|---|---|---|
| Sample: | Full | 1940 | 1940 | 1940 | 1940 | Full | Full |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| $\ln(\frac{S}{U})_{1980}$ | 0.153 | 0.142 | 0.129 | 0.100 | | 0.113 | 0.182 |
| | (0.0115) | (0.0129) | (0.0211) | (0.0212) | | (0.0219) | (0.0787) |
| $\ln(\frac{S}{U})_{1940}$ | | | | | 0.104 | | |
| | | | | | (0.0214) | | |
| $\Delta\ln(\frac{S}{U})_{1980-1940}$ | | | | | 0.0812 | | |
| | | | | | (0.0233) | | |
| R-squared | 0.448 | 0.454 | 0.537 | 0.662 | 0.668 | 0.602 | 0.582 |
| rMSE | 0.0480 | 0.0414 | 0.0398 | 0.0347 | 0.0345 | 0.0425 | 0.0436 |
| Test $H_0$: = coeffs | | | | | 2.141 | | |
| p-value | | | | | 0.146 | | |
| Controls | | | | | | | |
| Industry | | | yes | yes | yes | yes | yes |
| City | | | | yes | yes | yes | yes |
| Observations | 217 | 151 | 151 | 151 | 151 | 217 | 217 |

Heteroskedasticity-robust standard errors in parentheses. The dependent variable in all columns is the average number of personal computers per employee in the metro area, regression-adjusted for plant employment and three-digit SIC industry. Data source: Harte-Hanks. All other variables from U.S. censuses of the relevant year (aggregated to the metro area level). The variable $\ln(\frac{S}{U})$ corresponds to the log of the ratio of college equivalent workers – those with at least a four year degree plus half of those with some college – to non-college equivalent workers in each year. Industry controls are the shares of an area's 1980 employment in agriculture and mining, construction, non-durable manufacturing, durable manufacturing, transportation and utilities, wholesale, retail, finance and real estate, business and repair services, personal services, entertainment, and professional services (public sector share excluded). The city level controls are the log of the labor force, the unemployment rate, and the shares of city's working-age population which is female, African-American, and Mexican-born in 1980. All regressions weighted by $(\frac{1}{N_{12,1980}}+\frac{1}{N_{16,1980}}+\frac{1}{N_{12,2000}}+\frac{1}{N_{16,2000}})^{-0.5}$, where $N_{16}$ and $N_{12}$ represent the raw number of working-age census respondents in the area who report wages and have exactly 16 and 12 years, respectively, of education in the corresponding census. [1]The instrumental variable in column (7) is the share of 15-19-year-olds who report enrolling in school in the past year in 1880, constructed using a 10% sample of the 1880 U.S. Census (Ruggles et al., 2009). In the first stage, the coefficient on this variable is 0.45 with an F-statistic of 15.75.

Table 2: PCs per Worker, 2000, and 1980 Return to Skill

| Estimation: | OLS | OLS | OLS | OLS | IV[1] | IV[2] | OLS | IV[3] |
|---|---|---|---|---|---|---|---|---|
| Sample: | Full | 1940 | 1940 | 1940 | 1940 | 1940 | Full | Full |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $\ln(\frac{w^s}{w^u})_{1980}$ | -0.335 | -0.266 | -0.226 | -0.199 | -0.615 | -0.615 | -0.189 | -0.511 |
| | (0.101) | (0.131) | (0.0855) | (0.0743) | (0.156) | (0.155) | (0.0715) | (0.255) |
| R-squared | 0.069 | 0.045 | 0.628 | 0.628 | 0.556 | 0.556 | 0.562 | 0.520 |
| rMSE | 0.0624 | 0.0548 | 0.0363 | 0.0363 | 0.0397 | 0.0397 | 0.0446 | 0.0467 |
| over-ID | | | | | | 2.091 | | |
| p-value | | | | | | .148 | | |
| Controls | | | | | | | | |
| Industry | | | yes | yes | yes | yes | yes | yes |
| City | | | | yes | yes | yes | yes | yes |
| Observations | 217 | 151 | 151 | 151 | 151 | 151 | 217 | 217 |

Heteroskedasticity-robust standard errors in parentheses. The dependent variable in all columns is the average number of personal computers per employee in the metro area, regression-adjusted for plant employment and three-digit SIC industry. Data source: Harte-Hanks. All other variables from Census of Population of the relevant year. $\ln(\frac{w^s}{w^u})_{1980}$ represents the adjusted log hourly wage gap between those with exactly 16 and exactly 12 years of schooling. See Appendix B for details of the adjustment, which controls for a quartic in potential experience, and dummies for four age categories, white, married, female, foreign-born, being born after 1950 and its interaction with schooling. Industry controls are the shares of an area's 1980 employment in agriculture and mining, construction, non-durable manufacturing, durable manufacturing, transportation and utilities, wholesale, retail, finance and real estate, business and repair services, personal services, entertainment, and professional services (public sector share excluded). The city level controls are the log of the labor force, the unemployment rate, and the shares of city's working-age population which is female, African-American, and Mexican-born in 1980. All regressions weighted by $(\frac{1}{N_{12,1980}} + \frac{1}{N_{16,1980}} + \frac{1}{N_{12,2000}} + \frac{1}{N_{16,2000}})^{-0.5}$, where $N_{16}$ and $N_{12}$ represent the raw number of working-age census respondents in the area who report wages and have exactly 16 and 12 years, respectively, of education in the corresponding census. Instruments: [1]the log of the ratio of college equivalent workers – those with at least a four year degree plus half of those with some college – to non-college equivalent workers in 1980 (first stage F-stat on instrument = 50.11); [2]the log of the ratio of college equivalents to non-college equivalents in 1940 and the change in this log ratio from 1940-1980 (first stage F-stat on instruments = 31.29); [3]the share of 15-19-year-olds enrolled in school in the past year, constructed using a 10% sample of the 1880 U.S. Census (Ruggles et al., 2009) (first stage F-stat=20.61).

Table 3a: Changes in the Return to Skill and Initial Education

| 1980-End Year: | 2000 | 2000 | 2000 | 2000 | 2000 | 2000 | 2000 | 2008 | 2008 |
|---|---|---|---|---|---|---|---|---|---|
| Estimation: | OLS | OLS | OLS | OLS | OLS | OLS | IV[1] | OLS | OLS |
| Sample: | Full | 1940 | 1940 | 1940 | 1940 | Full | Full | 1940 | Full |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| $\ln(\frac{S}{U})_{1980}$ | 0.0707 | 0.0742 | 0.0976 | 0.0998 | | 0.106 | 0.209 | 0.172 | 0.170 |
| | (0.0158) | (0.0186) | (0.0324) | (0.0324) | | (0.0267) | (0.0798) | (0.0338) | (0.0275) |
| $\ln(\frac{S}{U})_{1940}$ | | | | | 0.101 | | | | |
| | | | | | (0.0323) | | | | |
| $\Delta\ln(\frac{S}{U})_{1980-1940}$ | | | | | 0.0947 | | | | |
| | | | | | (0.0347) | | | | |
| R-squared | 0.147 | 0.185 | 0.557 | 0.561 | 0.561 | 0.469 | 0.398 | 0.539 | 0.464 |
| rMSE | 0.0482 | 0.0416 | 0.0326 | 0.0326 | 0.0327 | 0.0398 | 0.0424 | 0.0411 | 0.0473 |
| $H_0$: = coefficients | | | | | 0.250 | | | | |
| *p-value* | | | | | 0.618 | | | | |
| Controls | | | | | | | | | |
| Industry | | | yes | yes | yes | yes | yes | yes | yes |
| City | | | yes | yes | yes | yes | yes | yes | yes |
| Endog. area? | | | | yes | yes | yes | yes | yes | yes |
| Observations | 217 | 151 | 151 | 151 | 151 | 217 | 217 | 151 | 217 |

Heteroskedasticity-robust standard errors in parentheses. Data sources: various Censuses of Population and the 2006-2008 American Community Survey (Ruggles et al., 2009). Dependent variable is the adjusted log hourly wage gap between those with exactly 16 and exactly 12 years of schooling between 1980 and 2000 (columns 1-7) and 1980 and 2008 (columns 8-9). The adjustment includes controls for a quartic in potential experience, and dummies for four age categories, white, married, female, foreign-born, being born after 1950 and its interaction with schooling. Columns with the "endogenous area" control also include a quadratic in the estimated probability of being located in the area, conditional on place of birth (state or foreign region), age, white, married, and female. See Appendix B for details. Industry controls are the shares of an area's 1980 employment in agriculture and mining, construction, non-durable manufacturing, durable manufacturing, transportation and utilities, wholesale, retail, finance and real estate, business and repair services, personal services, entertainment, and professional services (public sector share excluded). The city level controls are the log of the labor force, the unemployment rate, and the shares of city's working-age population which is female, African-American, and Mexican-born in 1980. All regressions weighted by $(\frac{1}{N_{12,1980}} + \frac{1}{N_{16,1980}} + \frac{1}{N_{12,2000}} + \frac{1}{N_{16,2000}})^{-0.5}$, where $N_{16}$ and $N_{12}$ represent the raw number of working-age census respondents in the area who report wages and have exactly 16 and 12 years, respectively, of education in the corresponding census. The independent variable $\ln(\frac{S}{U})$ corresponds to the log of the ratio of college equivalent workers – those with at least a four year degree plus half of those with some college – to non-college equivalent workers in each year.
[1]The instrumental variable is the share of 15-19-year-olds who report enrolling in school in the past year in 1880, constructed using a 10% sample of the 1880 U.S. Census (Ruggles et al., 2009). The first stage F-stat is 15.75.

Table 3b: Changes in the Return to Skill Before and After 1980, with Alternative Adjustments

| Estimation: | OLS | OLS | OLS | IV[1] |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | *A. Change in returns, 1940/50-1980* | | | |
| $\ln(\frac{S}{U})_{1980}$ | -0.00496 | -0.0357 | | -0.116 |
| | (0.0269) | (0.0584) | | (0.288) |
| $\ln(\frac{S}{U})_{1940}$ | | | -0.0296 | |
| | | | (0.0575) | |
| $\Delta \ln(\frac{S}{U})_{1980-1940}$ | | | -0.0700 | |
| | | | (0.0786) | |
| R-squared | 0.000 | 0.133 | 0.140 | 0.123 |
| rMSE | 0.0979 | 0.0969 | 0.0969 | 0.0975 |
| | *B. Change in returns, 1980-2000* | | | |
| $\ln(\frac{S}{U})_{1980}$ | 0.0525 | 0.0931 | | 0.158 |
| | (0.0205) | (0.0349) | | (0.0958) |
| $\ln(\frac{S}{U})_{1940}$ | | | 0.0934 | |
| | | | (0.0350) | |
| $\Delta \ln(\frac{S}{U})_{1980-1940}$ | | | 0.0918 | |
| | | | (0.0370) | |
| R-squared | 0.088 | 0.514 | 0.514 | 0.484 |
| rMSE | 0.0452 | 0.0350 | 0.0352 | 0.0361 |
| Controls | | | | |
| Industry | | yes | yes | yes |
| City | | yes | yes | yes |
| Observations | 151 | 151 | 151 | 151 |

Heteroskedasticity-robust standard errors in parentheses. Data sources for dependent and independent variables: various Censuses of Population. The dependent variable is a metro-specific return to college equivalent status (equal to one for those with at least four, 0.5 for those with fewer than four but at least one year of college, zero otherwise), adjusted for (non-metro-specific) years of education, a quartic in potential experience, and dummies for immigrant and female, in the sample with at least 11 years of education. The 1940 and 1950 data were combined and also include a year dummy control. The independent variable $\ln(\frac{S}{U})$ corresponds to the log of the ratio of college equivalent workers-to non-college equivalent workers in each year. Industry controls and city controls are as in previous tables. All regressions weighted by $(\frac{1}{N_{12,1980}} + \frac{1}{N_{16,1980}} + \frac{1}{N_{12,2000}} + \frac{1}{N_{16,2000}})^{-0.5}$, where $N_{16}$ and $N_{12}$ represent the raw number of working-age census respondents in the area who report wages and have exactly 16 and 12 years, respectively, of education in the corresponding census. [1]The instrumental variable in column (4) is the share of 15-19-year-olds who report enrolling in school in the past year in 1880, constructed using a 10% sample of the 1880 U.S. Census (Ruggles et al., 2009). The first stage F-stat is 8.07.

Table 4: Return to Skill and Skill Supply: 1940/50, 1980, 2000, and 2006-8

| Estimation: | OLS | OLS | OLS | OLS | OLS | OLS | OLS | OLS |
|---|---|---|---|---|---|---|---|---|
| Returns in Year: | 1940/50[1] | 1980 | 2000 | 2006-8 | 1940/50[1] | 1980 | 2000 | 2006-8 |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $\ln(\frac{S}{U})_{1940}$ | -0.0464 | | | | -0.0886 | | | |
| | (0.0288) | | | | (0.0379) | | | |
| $\ln(\frac{S}{U})_{1980}$ | | -0.0638 | | | | -0.169 | | |
| | | (0.0131) | | | | (0.0195) | | |
| $\ln(\frac{S}{U})_{2000}$ | | | 0.00769 | | | | -0.0331 | |
| | | | (0.0170) | | | | (0.0269) | |
| $\ln(\frac{S}{U})_{2006-8}$ | | | | 0.0242 | | | | 0.0350 |
| | | | | (0.0177) | | | | (0.0263) |
| R-squared | 0.023 | 0.128 | 0.001 | 0.012 | 0.184 | 0.526 | 0.528 | 0.562 |
| rMSE | 0.0951 | 0.0472 | 0.0570 | 0.0589 | 0.0924 | 0.0361 | 0.0408 | 0.0408 |
| Controls | | | | | | | | |
| Industry | | | | | yes | yes | yes | yes |
| City | | | | | yes | yes | yes | yes |
| Endog Area? | | | | | | yes | yes | yes |
| Observations | 151 | 217 | 217 | 217 | 151 | 217 | 217 | 217 |

Heteroskedasticity-robust standard errors in parentheses. Data sources for dependent and independent variables: various Censuses of Population. [1]The dependent variable in columns (1) and (5) is a metro-specific return to college equivalent status (equal to one for those with at least four, 0.5 for those with fewer than four but at least one year of college, zero otherwise), adjusted for (non-metro-specific) years of education, a quartic in potential experience, and dummies for immigrant, female, and year, in the stacked 1940 and 1950 observations with at least 11 years of education. In all other columns, the dependent variable is the log hourly wage gap between those with exactly 16 and exactly 12 years of schooling, regression adjusted for a quartic in potential experience and dummies for four age categories, white, married, female, foreign-born, being born after 1950 and its interaction with schooling. Columns with the "endogenous area" control also include a quadratic in the estimated probability of being located in the area, conditional on place of birth (state or foreign region), age, white, married, and female. See Appendix B for details. The independent variable $\ln(\frac{S}{U})$ corresponds to the log of the ratio of college equivalent workers-to non-college equivalent workers in each year. Industry controls and city controls are as in previous tables. All regressions weighted by $(\frac{1}{N_{12,1980}} + \frac{1}{N_{16,1980}} + \frac{1}{N_{12,2000}} + \frac{1}{N_{16,2000}})^{-0.5}$, where $N_{16}$ and $N_{12}$ represent the raw number of working-age census respondents in the area who report wages and have exactly 16 and 12 years, respectively, of education in the corresponding census.

# A Appendix A: Proofs

The competitive equilibrium of an economy faced with the choice between two techniques of production will solve the following program:

$$\max_{\gamma,\rho,K,PC} K^{1-\alpha}[a(\gamma S)^{\sigma}+(1-a)(\rho U)^{\sigma}]^{\frac{\alpha}{\sigma}}+PC^{1-\alpha}[b((1-\gamma)S)^{\sigma}+(1-b)((1-\rho)U)^{\sigma}]^{\frac{\alpha}{\sigma}}-r^{K}K-r^{PC}PC$$

subject to $0 \leq \gamma \leq 1$, $0 \leq \rho \leq 1$, $0 \leq K$ and $0 \leq PC$

Recall that each local economy takes the price of the two types of capital as given. (Note: In this appendix the subscript $i$ on $S_i$ and $U_i$ are dropped for simplicity) The main first order conditions associated with this this problem are:

$$\begin{aligned}
&(1-\alpha)K^{-\alpha}[a(\gamma S)^{\sigma}+(1-a)(\rho U)^{\sigma}]^{\frac{\alpha}{\sigma}} \leq r^{K}\\
&(1-\alpha)PC^{-\alpha}[b((1-\gamma)S)^{\sigma}+(1-b)((1-\rho)U)^{\sigma}]^{\frac{\alpha}{\sigma}} \leq r^{PC}
\end{aligned}$$

$$\begin{aligned}
&\text{If} \quad 0<\gamma<1\\
&K^{1-\alpha}[a(\gamma S)^{\sigma}+(1-a)(\rho U)^{\sigma}]^{\frac{\alpha}{\sigma}-1}a\gamma^{\sigma-1}\\
=&PC^{1-\alpha}[b((1-\gamma)S)^{\sigma}+(1-b)((1-\rho)U)^{\sigma}]^{\frac{\alpha}{\sigma}-1}b(1-\gamma)^{\sigma-1}\\
&\text{and} \quad \text{if} \quad 0<\rho<1\\
&K^{1-\alpha}[a(\gamma S)^{\sigma}+(1-a)(\rho U)^{\sigma}]^{\frac{\alpha}{\sigma}-1}(1-a)\rho^{\sigma-1}\\
=&PC^{1-\alpha}[b((1-\gamma)S)^{\sigma}+(1-b)((1-\rho)U)^{\sigma}]^{\frac{\alpha}{\sigma}-1}(1-b)(1-\rho)^{\sigma-1}
\end{aligned}$$

In order to characterize the solution to the problem, let us define $\phi_L$ and $\phi_L$ implicitly by the following two equations

$$(\frac{1}{r^K})^{\frac{\alpha}{1-\alpha}}a[a+(1-a)(\frac{1}{\phi^L})^{\sigma}]^{\frac{1-\sigma}{\sigma}}=(\frac{1}{r^{PC}})^{\frac{\alpha}{1-\alpha}}b[b+(1-b)(\frac{1}{\phi^H})^{\sigma}]^{\frac{1-\sigma}{\sigma}}$$

and

$$(\frac{1}{r^K})^{\frac{\alpha}{1-\alpha}}(1-a)[a(\phi^L)^\sigma + (1-a)]^{\frac{1-\sigma}{\sigma}} = (\frac{1}{r^{PC}})^{\frac{\alpha}{1-\alpha}}(1-b)[b(\phi^H)^\sigma + (1-b)]^{\frac{1-\sigma}{\sigma}}$$

Under Assumption 2, the following values for $\gamma$, $\rho$ , $K$ and $PC$ satisfy the first order conditions, including the associated complementarity slackness conditions, and hence constitute a solution to the problem.

If $\frac{S}{U} < \phi_L$, then $\gamma = 1$, $\rho = 1$ and $PC = 0$ and $K = (1-\alpha)^{\frac{1}{\alpha}}(r^K)^{\frac{-1}{\alpha}}[a(S)^\sigma + (1-a)(U)^\sigma]^{\frac{1}{\sigma}}$

If $\phi_L \leq \frac{S}{U} \leq \phi_H$, then $\gamma = \frac{\frac{\phi_H U}{S} - 1}{\frac{\phi_H}{\phi_L} - 1}$, $\rho = \frac{\frac{\phi_H}{\phi_L} - \frac{S}{\phi_L U}}{\frac{\phi_H}{\phi_L} - 1}$, $K = (1-\alpha)^{\frac{1}{\alpha}} r^{K^{\frac{-1}{\alpha}}}[a(\gamma S)^\sigma + (1-a)(\rho U)^\sigma]^{\frac{1}{\sigma}}$ and $PC = (1-\alpha)^{\frac{1}{\alpha}} r^{PC^{\frac{-1}{\alpha}}}[b((1-\gamma)S)^\sigma + (1-b)((1-\rho)U)^\sigma]^{\frac{1}{\sigma}}$

If $\frac{S}{U} > \phi_H$, then $\gamma = 0$, $\rho = 0$ and $K = 0$ and $PC = (1-\alpha)^{\frac{1}{\alpha}}(r^{PC})^{\frac{-1}{\alpha}}[b(S)^\sigma + (1-b)(U)^\sigma]^{\frac{1}{\sigma}}$

This characterization of the solution to the optimization problem will be used in the proofs of the propositions and corollaries.

**Proof of Proposition 1:** From the above solution to the maximization problem, we know that PCs per worker will be given by:

$If \quad \frac{S}{U} > \phi_H, \quad \frac{PC}{S+U} = (1-\alpha)^{\frac{1}{\alpha}}(r^{PC})^{\frac{-1}{\alpha}}[b(\frac{S}{S+U})^\sigma + (1-b)(\frac{U}{S+U})^\sigma]^{\frac{1}{\sigma}}$

which under Assumption 1 is an increasing function of $\frac{S}{U}$

$If \quad \phi_L \leq \frac{S}{U} \leq \phi_H, \quad \frac{PC}{S+U} = (1-\alpha)^{\frac{1}{\alpha}} r^{PC^{\frac{-1}{\alpha}}}[b((1-\gamma)\frac{S}{S+U})^\sigma + (1-b)((1-\rho)\frac{U}{S+U})^\sigma]^{\frac{1}{\sigma}}$

where $\gamma = \frac{\frac{\phi_H U}{S} - 1}{\frac{\phi_H}{\phi_L} - 1}$ and $\rho = \frac{\frac{\phi_H}{\phi_L} - \frac{S}{\phi_L U}}{\frac{\phi_H}{\phi_L} - 1}$

which again is an increasing function of $\frac{S}{U}$

and finally, if $\frac{S}{U} < \phi_L$, then $\frac{PC}{S+U} = 0$.

Hence, PC per worker is a weakly increasing function of $\frac{S}{U}$.

**Proof of Corollary 1:** Since by Proposition 1, PCs per worker is an increasing function of a $\frac{S}{U}$ and the initial ratio of skilled to unskilled wages is given by:

$$\frac{w^S}{w^U} = \frac{aS^{\sigma-1}}{(1-a)U^{\sigma-1}}$$

it follows that PCs per worker is an increasing function of the locality's initial ratio of skilled to unskilled wages

**Proof of Proposition 2:** Before the arrival of the new technology, the relationship between returns to skill and the supply of skill is given by:

$$\ln(\frac{w^S}{w^U}) = \ln(\frac{aS^{\sigma-1}}{(1-a)U^{\sigma-1}})$$

After the arrival of the new technology, the relationship is given by:

$$\begin{array}{llll}
\ln(\frac{w^S}{w^U}) = & \ln(\frac{aS^{\sigma-1}}{(1-a)U^{\sigma-1}}) & if & \frac{S}{U} \leq \phi^L \\
\ln(\frac{w^S}{w^U}) = & \ln(\frac{a(\phi_L)^{\sigma-1}}{(1-a)}) = \ln(\frac{b(\phi_H)^{\sigma-1}}{(1-b)}) & if & \phi^L < \frac{S}{U} \leq \phi^H \\
\ln(\frac{w^S}{w^U}) = & \ln(\frac{bS^{\sigma-1}}{(1-b)U^{\sigma-1}}) & if & \phi^H < \frac{S}{U}
\end{array}$$

Hence, the relationship between the initial supply of skill and the change in the returns to skill is given by:

$$\begin{array}{llll}
\Delta \ln \frac{w^S}{w^U} = & 0 & if & \frac{S}{U} \leq \phi^L \\
\Delta \ln \frac{w^S}{w^U} = & (1-\sigma)[\log \frac{S}{U} - \log \phi^L] & if & \phi^L < \frac{S}{U} \leq \phi^H \\
\Delta \ln \frac{w^S}{w^U} = & (1-\sigma)[\log \phi^H - \log \phi^L] & if & \phi^H < \frac{S}{U}
\end{array}$$

This implies that the change in returns to skill is positively related to the initial supply of skill.

**Proof of Corollary 2:** Corollary 2 follows directly by combining Propositions 1 and 2.

**Proof of Proposition 3:** After the arrival of the skill-biased technology, the relationship between the supply of skill and the return to skill is given by

$$
\begin{aligned}
\ln(\frac{w^S}{w^U}) &= \ln(\frac{aS^{\sigma-1}}{(1-a)U^{\sigma-1}}) && if && \frac{S}{U} \leq \phi^L \\
\ln(\frac{w^S}{w^U}) &= \ln(\frac{a(\phi_L)^{\sigma-1}}{(1-a)}) = \ln(\frac{b(\phi_H)^{\sigma-1}}{(1-b)}) && if && \phi^L < \frac{S}{U} \leq \phi^H \\
\ln(\frac{w^S}{w^U}) &= \ln(\frac{bS^{\sigma-1}}{(1-b)U^{\sigma-1}}) && if && \phi^H < \frac{S}{U}
\end{aligned}
$$

Since this function exhibits a negative relationship between the supply of skill and the return to skill, the arrival of the skill-biased technology cannot lead to a positive association between the return to skill and the supply of skill.

**Proof of Corollary 3:** Corollary 3 follows directly from Proposition 3 and 1.

**Proof of Proposition 4:** Before the arrival of the new technology, $\frac{\partial \ln \frac{W^S}{W^U}}{\partial \ln \frac{S}{u}}$ is equal to $1-\sigma$. After the arrival of the new technology it is:

$$
\begin{aligned}
\frac{\partial \ln \frac{W^S}{W^U}}{\partial \ln \frac{S}{u}} &= 1-\sigma && if && \frac{S}{U} \leq \phi^L \\
\frac{\partial \ln \frac{W^S}{W^U}}{\partial \ln \frac{S}{u}} &= 0 && if && \phi^L < \frac{S}{U} \leq \phi^H \\
\frac{\partial \ln \frac{W^S}{W^U}}{\partial \ln \frac{S}{u}} &= 1-\sigma && if && \phi^H < \frac{S}{U}
\end{aligned}
$$

Hence, the effect of an increase in supply of skill on the returns to skill is (weakly) smaller after the arrival of the new technology. In particular, the effect is zero when the economy is in the technology transition zone $\phi^L \leq \frac{S}{U} \leq \phi^H$.

**Proof of Proposition 5:**

Let us begin by proving point (2) of the proposition by contradiction. We now reintroduce subscripts on $S$ and $U$ since we need to compare different localities. As shown in the text, endogenous mobility implies,

$$
\log(\frac{S_i}{U_i}) - \log(\frac{S_j}{U_j}) = \frac{\psi}{\psi-1}[\log(\frac{d_j}{d_i}) + (\frac{1+\upsilon}{\upsilon})(\log(\frac{w_i^s}{w_i^u}) - \log(\frac{w_j^s}{w_j^u}))]
$$

which means that if locality $i$ has a greater increase in returns to skill than locality $j$, and it has a lower initial supply of skill, it has a greater (log) increase in the supply of skill. Now suppose this is the case and locality $i$ has a initial skill level that is lower than locality $j$, but has a higher (log) increase in the supply of skill. This would imply that an increase in supply has a greater effect after the arrival of the new technology than before, which would contradict Proposition 4. Hence, point (2) of the proposition must hold.

From the proof of point (2) of the proposition, we know that a city with an initially higher fraction of skilled to unskilled workers will maintain its higher skill ratio after the arrival of the new technology and hence point (1) of Proposition 5 follows from the relationship between PCs per workers and skill given in Proposition 1. Finally point (3) of Proposition 5 follows directly from Proposition 3 and Corollary 3.

## A.1 Decomposing Pre-PC Skill Variation

To simplify the notation, drop the "i" subscript for localities and define $X_t \equiv \ln(\frac{S}{U})_{i,t}$, and let $\Delta Y$ represent some post-1980 outcome (adoption of PCs or changes in the return to skill). The data generating process is assumed to be:

$$\Delta Y = \gamma_1 X_{1980} + \gamma_2 Z_{1980} + \epsilon_t$$

with the unobserved "innovative city" factor, $Z$, related to skill by a linear relationship:[73]

$$X_t = \psi Z_t + \eta_t$$

$\epsilon_t$ is independent of everything, and $\eta_t$ is independent of $Z_t$ in all periods. The problem arises when $\gamma_2 \neq 0$ and $\psi \neq 0$. Now, without much loss of generality (given we are looking over only two periods, 1940 and 1980), assume that the autocorrelation of the components of X are given by:

$$Z_t = \nu_t + \alpha \nu_{t-1}$$

[73]The constancy of $\psi$ (no "t" subscript) embodies the idea, for example, of constant capital-skill complementarity over the twentieth century (Goldin & Katz, 2008). This assumption is not made for our convenience, rather, it is one reason that our test could have low power. See below.

$$\eta_t = \xi_t + \rho\xi_{t-1}$$

where $\nu$ and $\xi$ are iid. This structure captures the idea that there can be autocorrelation in both $Z_t$ and $X_t$. Now substitute in the decomposition $X_{1980} \equiv \Delta X_{1940-80} + X_{1940}$:

$$\Delta Y = \gamma_{11}\Delta X_{1940-80} + \gamma_{12}X_{1940} + \gamma_2 Z_{1980} + \epsilon_t$$

(so $t = 1980$ and $t-1 = 1940$ and the true coefficients satisfy $\gamma_{11} = \gamma_{12} = \gamma_1$). For simplicity normalize $\psi = 1$. Under these assumptions, it can be shown that the probability limit of the difference between the OLS estimates of $\gamma_{11}$ and $\gamma_{12}$ is proportional to:[74]

$$\gamma_2((1+\rho^2)(1-\alpha+\alpha^2) - (1-\rho+\rho^2)(1+\alpha^2))\sigma_\nu^2\sigma_\xi^2$$

and hence testing whether $\gamma_{11} = \gamma_{12}$ is equivalent to testing whether the above expression is zero. As can be seen, this expression will be zero if either a) $\gamma_2 = 0$ (no omitted variable bias) or b) $\rho = \alpha$. If you take the latter to be a zero probability event, then we have a consistent test. However, if $\rho$ is close to $\gamma$, then it may not have much power. By looking at periods further apart, we increase the power.[75]

[74] In terms of X's and Z's the expression comes from multivariate OLS $\gamma_2[Cov(\Delta X, Z) - Cov(X_{t-1}, Z)][Cov(\Delta X, X_{t-1}) + Var(X_{t-1})]/[Var(\Delta X)Var(X_{t-1}) - Cov^2(\Delta X, X_{t-1})]$, the numerator of which reduces to the expression shown after substituting in the MA(1) processes for X and Z.

[75] Note that the expression is more complex if $\psi$ is changing over time, and this is another reason that $\hat{\gamma}_{11} - \hat{\gamma}_{12}$ is unlikely to be centered on zero in widely separated periods.

# B Appendix B: Wage Construction and Adjustments

All wage variables were constructed from public use micro data from the US Census of Population (various years) using employed workers between the ages of 16 and 65, with at least one year of "potential experience" (age - years of education - 6), positive hours worked, not living in group quarters, and (except in Table 3b) having either exactly 12 or 16 years of education. We also trimmed the wage sample to include only those with hourly earnings between \$2 and \$200 in 1999 dollars. To obtain regression-adjusted returns, the natural log of hourly wages was regressed on a college dummy interacted with a dummy for each metropolitan area, controlling for metropolitan area effects, a quartic in potential work experience, a female dummy, a foreign-born dummy, a dummy for being white, a dummy for being married, dummies for three age categories (30-39, 40-49, 50-65), a dummy for being born after 1950 (which is roughly when a trend break in aggregate skill supply occurs), and its interaction with a college graduate dummy. Adjusted returns come from the vector of coefficients on the college x area dummies.

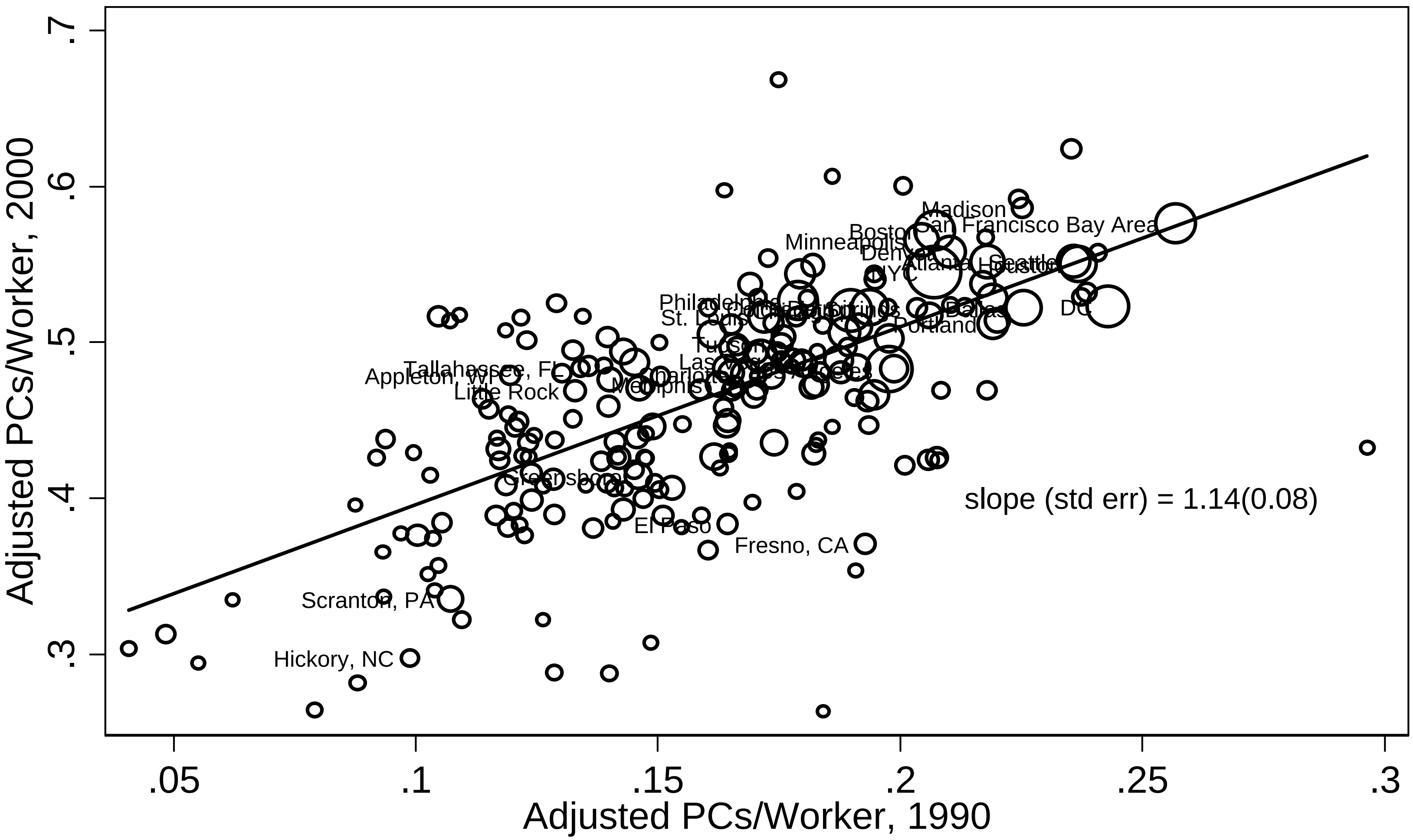
In the regressions labeled with "Engog. Area?" in the tables, wages were further adjusted for the selection of labor market using a procedure similar to the one outlined in Dahl (2002). (More details below.) In practice, this means adding a control for a polynomial in the probability of being a wage earner in a particular metropolitan area conditional on place of birth and other demographic characteristics. Dahl allowed for separate control functions for two types of residents: "stayers" - those living in their state of birth - and "movers." Like Beaudry, Green, and Sand (2007), on which our implementation is largely based, we add a third resident type, foreign-born, who were excluded from Dahl's analysis.

In our paper (in contrast to Dahl, 2002) metropolitan area, not state, is the unit of analysis. Metropolitan area of birth is not identified in the Census. To accommodate this, we follow Beaudry, Green, and Sand (2007) and define a "stayer" as any individual who lives in a metropolitan area which is at least partly in the individual's state of birth. For example, some counties in West Virginia are defined to be part of the Washington, DC metropolitan area. As a result, someone who was born in West Virginia and who lives anywhere in the DC area is considered a stayer.

Among the native-born (stayers and movers), an individual's probability of being located in the area was estimated as the share of people in the same demographic "cell" living in that metropolitan area. The demographic cells were defined by the interaction of two education categories (college, high school), four age categories (16-29, 30-39, 40-49, 50-65), two race categories (white/nonwhite), gender, a dummy for married, and state of birth. These shares were computed separately for movers and stayers. Among the foreign-born, the probability of living in a particular metropolitan area was allowed to vary across 18 region-of-origin groups: the 16 foreign regions used in Lewis (2003), plus the U.S. territory of Puerto Rico, plus all other foreign births (including other U.S. territories and U.S. natives born abroad of American parents). The immigrant portion of the selection correction thus corrects for endogeneity in a manner similar to the "supply/push"-type instruments used in research on the local labor market impact of immigration. Implicitly, the procedure treats Mexican immigration to border areas like Los Angeles as not very selected (the share of Mexicans living in LA is high) and treats Mexican immigration to Scranton, PA (where there are few Mexicans) as highly selected.

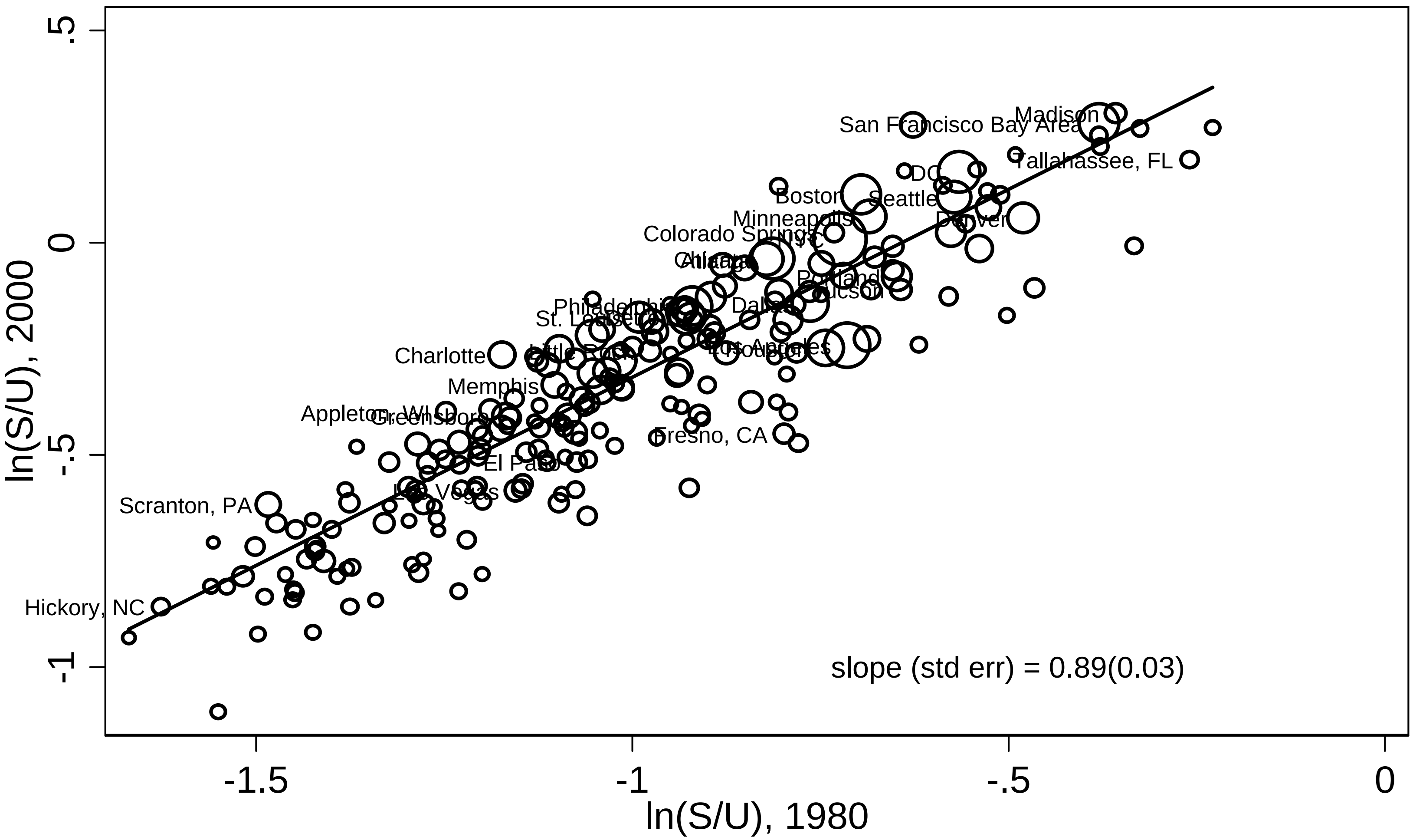
The adjustment of returns in "Endog. Area?" columns ultimately includes, in addition to all of the above controls, dummies for resident type (mover, stayer, or foreign-born), and quadratics for the probability of moving to a particular area for US movers (set to zero for non-movers and foreign-born), the probability of moving to the area for foreign-born (set to zero for all US born), probability of being a stayer for stayers (set to zero for foreign-born and movers), and the probability of being a stayer for movers (set to zero for foreign-born and stayers). The first three quadratics were statistically significantly related to wages, while the latter was not.

Figure 3: Adjusted PCs Per Worker
2000 vs. 1990



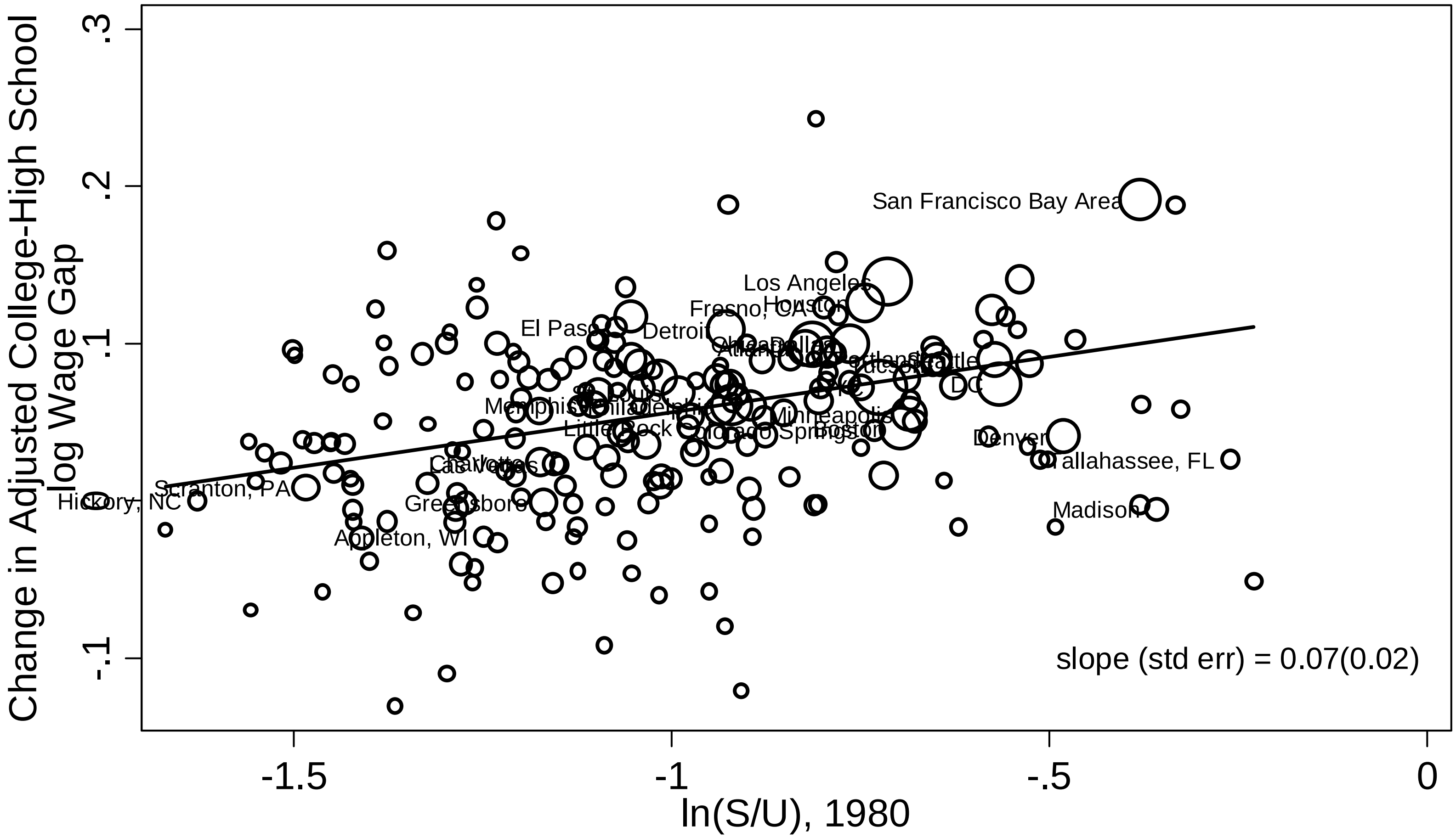

The size of the circles is proportional to the weights used in all of the regressions (see tables); weights used to fit the estimated line. Raw data, from Harte-Hanks, are adjusted for 3-digit SIC x 8 class size interactions, separately in 1990 and 2000, and the adjusted metro-area means are shown.

Figure 4: Log of College to Non-College Equivalents
2000 vs. 1980



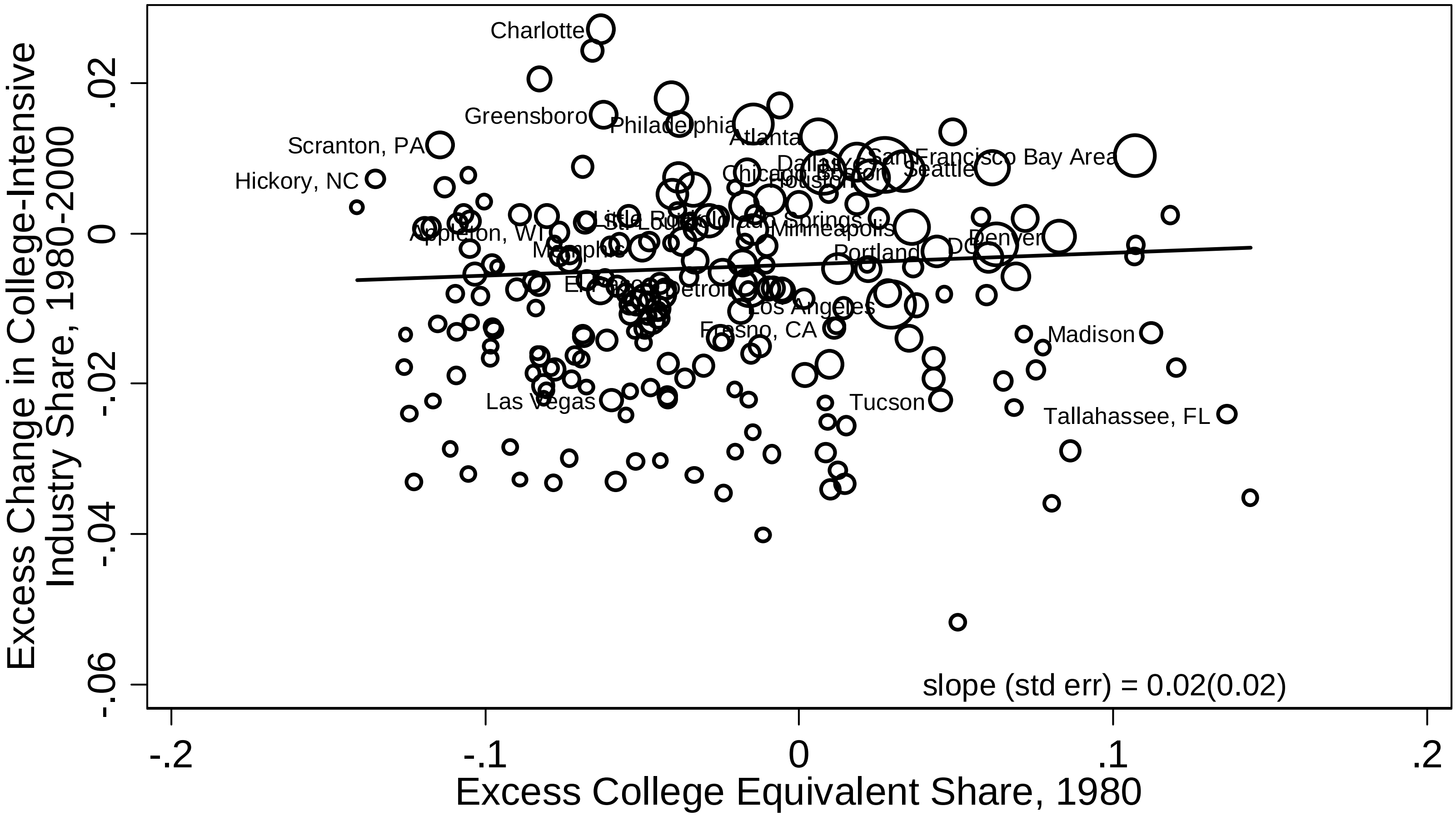

The size of the circles is proportional to the weights used in all of the regressions (see tables); weights used to fit the estimated line. College equivalents are defined as those with at least a four-year college degree plus 1/2 of those with some college. Raw data are from the 1980 and 2000 5% public-use Census of Population (PUMS).

Figure 5: Change in Adjusted College-High School log Wage Gap, 1980-2000
vs. log Ratio of College to Non-College Equivalents, 1980




The size of the circles is proportional to the weights used in all of the regressions (see tables); weights used to fit the estimated line. College high school-wage gap adjusted using methods described in Appendix B. College equivalents are defined as those with at least a four-year college degree plus 1/2 of those with some college. Raw data are from the 1980 and 2000 5% public-use Census of Population (PUMS).

Figure 6. College-Intensity Weighted Changes in Industry Mix vs. Initial College Equivalent Share




College equivalent share is the share of workers with a college degree plus 1/2 of those with some college education but no degree. Excess share is the college share in the area minus the college share in all cities. The y-axis shows how much excess demand for college labor was generated by changes in industry mix between 1980 and 2000 (calculation described in text). The size of the circles is proportional to the weights used in all of the regressions (see tables); weights used to fit the estimated line. Raw data from 1980 and 2000 5% public-use Census of Population (PUMS).